Physics

Student Textbook
Grade 10

Authors: Susan Gardner
Catherine Gaunt
Graham Bone

Advisers: Tilahun Tesfaye Deressu (PhD)
Endeshaw Bekele Buli

Evaluators: Yoseph Mihiret Mengistu
Gebremeskel Gebreegziabher
Yusuf Mohamed

Federal Democratic Republic of Ethiopia
Ministry of Education

PEARSON

Acknowledgments

The development, printing and distribution of this student textbook has been funded through the General Education Quality Improvement Project (GEQIP), which aims to improve the quality of education for Grades 1–12 students in government schools throughout Ethiopia.

The Federal Democratic Republic of Ethiopia received funding for GEQIP through credit/financing from the International Development Associations (IDA), the Fast Track Initiative Catalytic Fund (FTI CF) and other development partners – Finland, Italian Development Cooperation, the Netherlands and UK aid from the Department for International Development (DFID).
The Ministry of Education wishes to thank the many individuals, groups and other bodies involved – directly and indirectly – in publishing the textbook and accompanying teacher guide.

The publisher would like to thank the following for their kind permission to reproduce their photographs:

(Key: b-bottom; c-centre; l-left; r-right; t-top)

Alamy Images: David Ball 198r, Bruce Coleman Inc 55r, David J Green - electrical 71, Phil Degginger 127, Eureka 198l, Image Source 3, Ingram Publishing 57, mediacolor's 55l, Mario Moreno 37l, Jonathan Nightingale 37r, North Wind Picture Archives 6, prodme 219, World History Archive 53, 68, World History Archive 53, 68; **Corbis:** Bettmann 54, 62, Bettmann 54, 62, Steve Boyle / NewSport 35;
Getty Images: Juan Silva / The Image Bank 56; **Science Photo Library Ltd:** Professor Harold Edgerton 7, Kenneth Eward / Biografx 4, David Parker 231, Erich Schrempp 212t

Cover images: Front: **Alamy Images:** Bruce Coleman Inc br; **Getty Images:** Juan Silva / The Image Bank tr;
Science Photo Library Ltd: David Parker l

All other images © Pearson Education

Every effort has been made to trace the copyright holders and we apologise in advance for any unintentional omissions. We would be pleased to insert the appropriate acknowledgement in any subsequent edition of this publication.

First edition, 2002 (E.C.)
ISBN: 978-99944-2-018-6

Developed, Printed and distributed for the Federal Democratic Republic of Ethiopia, Ministry of Education by:
Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England

In collaboration with
Shama Books
P.O. Box 15
Addis Ababa
Ethiopia

Printed in Malaysia

Contents

Motion in 2D

Unit 1

Contents

Section	Learning competencies
1.1 Projectile motion (page 2)	• Define the term projectile (and provide several examples). • Explain the difference between 1D and 2D motion. • Correctly use the terms angle of elevation and angle of depression, and explain the importance of the angle when it comes to launching projectiles. • Explain the effect gravity has on the motion of an object. • Describe what happens to the horizontal and vertical velocities of a projectile and the important characteristics of its flight. • Demonstrate how to use the equations for uniform acceleration and to apply these to projectile motion. • Define the term centre of mass. • Conduct simple experiments to determine the centre of mass of 2D objects. • List the characteristics of uniform circular motion. • Describe the relationships between radius, mass, forces and velocity for an object following a circular path.
1.2 Rotational kinematics (page 24)	• Define the terms angular and tangential displacement, and angular and tangential velocity. • Express angles in terms of revolutions, radians and degrees. • Define the term angular acceleration, and list its key characteristics. • Identify the SI unit of angular velocity and angular acceleration. • Explain the relationships between angular displacement, tangential displacement, angular velocity, tangential velocity and angular acceleration. • Demonstrate how to use the equations of constant angular acceleration and compare them with equations of constant acceleration.
1.3 Rotational dynamics (page 30)	• Define the moment of inertia of a point mass. • Define rotational kinetic energy of a body. • Solve simple problems relating to moment of inertia and rotational kinetic energy. • Define the term torque. • Identify the SI unit of torque, N m, which is not the same as the joule. • Express torque in terms of moment of inertia and angular acceleration. • Derive an expression for the work done by the torque. • Use the formula $W = \tau\theta$ to solve problems related to work done by torque.

Contents

Section	Learning competencies
1.3 Rotational dynamics (page 30) (*continued*)	• Define the angular momentum of a particle of mass m and write its SI unit. • State the law of conservation of angular momentum. • Solve problems using the law of conservation of angular momentum. • State the first and second conditions of equilibrium. • Solve problems related to conditions of equilibrium. • Define the term centre of mass (centre of gravity) of a solid body. • Determine the centre of gravity using a plumb-line method. • Define the terms stable, unstable and neutral equilibrium.
1.4 Newton's law of universal gravitation (page 40)	• State Newton's law of universal gravitation. • Determine the magnitude of the force of attraction between two masses separated by a distance r. • Calculate the value of g at any distance above the surface of the Earth. • State Kepler's law of planetary motion. • Use Kepler's law of planetary motion to determine the period of any planet. • Differentiate between orbital and escape velocity of a satellite. • Determine the period of a satellite around a planet. • Calculate the orbital and escape velocity of a satellite. • Describe the period, position and function of a geostationary satellite.

You may have already studied motion in one dimension (1D), including the equation of uniform acceleration. This unit takes motion further, looking into motion in two dimensions. This includes projectiles and circular motion.

When you catch a ball your brain is completing a series of complex calculations relating to the path followed by the ball and the time it takes to reach you. This is hard-wired into our brains from the days when we used to hunt and most of our food lived in trees. A detailed understanding of two-dimensional (2D) motion is essential for physicists, as it enables them to complete the calculations required to design objects from complex rockets to roads around cities.

1.1 Projectile motion

By the end of this section you should be able to:

- Define the term projectile (and provide several examples).
- Explain the difference between 1D and 2D motion.
- Correctly use the terms angle of elevation and angle of depression, and explain the importance of the angle when it comes to launching projectiles.

- Explain the effect gravity has on the motion of an object.
- Describe what happens to the horizontal and vertical velocities of a projectile and the important characteristics of its flight.
- Demonstrate how to use the equations for uniform acceleration and to apply these to projectile motion.
- Define the term centre of mass.
- Conduct simple experiments to determine the centre of mass of 2D objects.
- List the characteristics of uniform circular motion.
- Describe the relationships between radius, mass, forces and velocity for an object following a circular path.

What are projectiles?

A **projectile** is any object moving through the air without an engine or other motive force. This means is it not restricted to cannonballs or bullets. When you throw a stone, toss a cricket ball, or kick a soccer ball, they are classed as projectiles as they fly through the air.

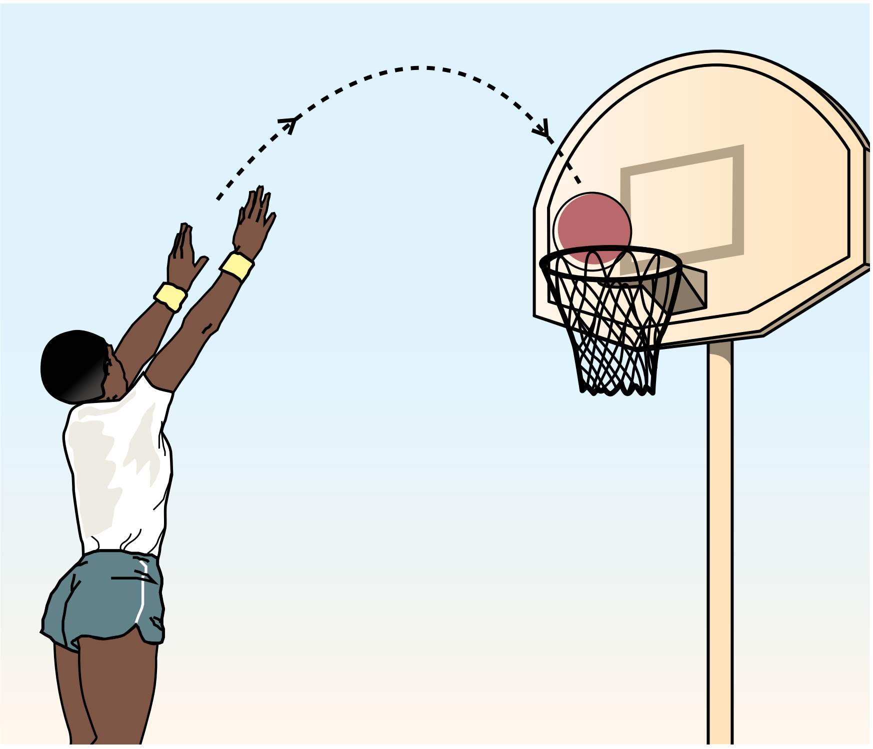


Figure 1.1 *Typical example of a projectile.*

Projectile motion is more complex than 1D motion. It involves motion in two directions, the vertical and the horizontal.

For simplicity we look at the horizontal **velocity** and vertical velocity as separate components of the velocity. These can be treated individually.

KEY WORDS

projectile *any object propelled through space by the exertion of a force which ceases after launch*

velocity *the rate of change of position of a body*

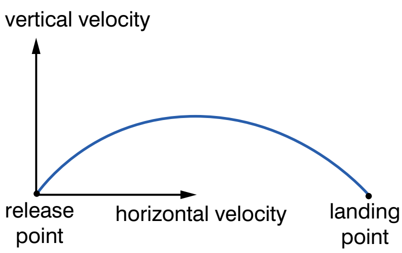


Figure 1.2 *Another projectile.*

KEY WORDS

resolving *splitting a vector into vertical and horizontal components. These components have the same effect as the original vector.*

Activity 1.1: Naming some projectiles

Spend 30 seconds listing as many projectiles as you can think of. Compare your list with a partner.

Resolving velocity

Resolving means splitting one vector into two component vectors (usually one horizontal and one vertical). These components have the same effect as the original vector.

An example can be seen below. The 60 m/s velocity can be resolved into two component vectors that have the same effect when combined:

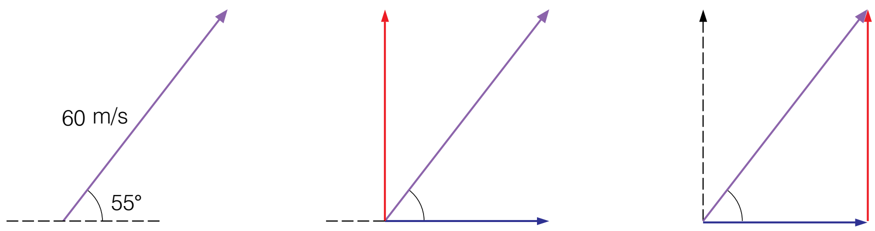


Figure 1.3 *Velocity components are shown in red and blue.*

This allows us to calculate the path of the projectile including its maximum vertical displacement (maximum height), maximum horizontal displacement (range) and flight time.

Horizontal motion

If we ignore air resistance, then there are no horizontal forces acting on the projectile as it flies through the air. This means there is no acceleration, so the velocity stays the same horizontally. This is a perfectly valid assumption for many projectiles.

We can apply the following equation:

displacement = average velocity × time taken

This becomes:

horizontal displacement = horizontal velocity × flight time

So in order to determine the range of the projectile we would use the total flight time. For example, to find the horizontal displacement after 5.0 s we would use a time of 5.0 seconds.

total horizontal displacement = horizontal velocity × total flight time

horizontal displacement after 5.0 s = velocity × 5.0 s

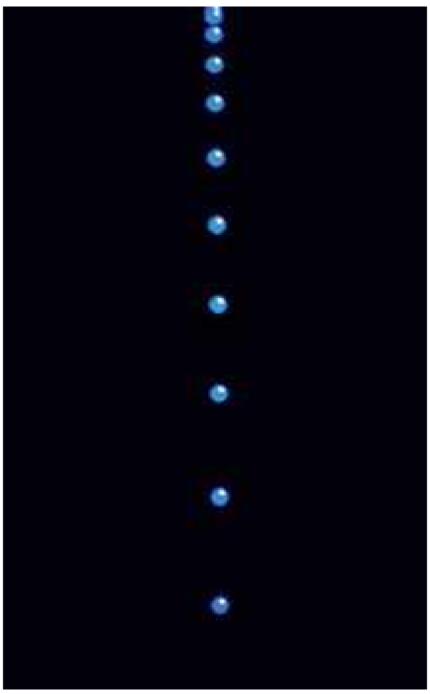


Figure 1.4 *Accelerating under gravity.*

Discussion activity

Under what circumstances might it be inappropriate to assume the horizontal velocity of a projectile remains constant?

Vertical motion

The vertical velocity does change. This is because the projectile accelerates under gravity as it moves.

Figure 1.4 shows a ball falling through the air. The images were taken at regular time intervals and you can see the displacement between each shot increases. This is because the ball is moving faster and faster.

Ignoring air resistance once again, gravity causes all objects to accelerate at 9.81 m/s^2.

Figure 1.5 shows two motion graphs for the falling ball. Looking at the displacement–time graph it is evident the gradient is increasing. This is because the object is moving faster as it falls.

This can be seen on the velocity–time graph. A constant gradient indicates a constant acceleration (in this case, 9.81 m/s^2).

The vertical motion of the projectile is an example of uniformly accelerated motion. This means we can use the equations for uniform acceleration:

1. $v = u + at$
2. $s = \frac{1}{2}(u + v)\,t$
3. $s = ut + \frac{1}{2}at^2$
4. $v^2 = u^2 + 2as$
5. $s = vt - \frac{1}{2}at^2$

where:

s = displacement

v = final velocity

u = initial velocity

a = acceleration (in this case, 9.81 m/s^2)

t = time

These can be used to determine the time it takes for a projectile to hit the ground.

Worked example 1.1

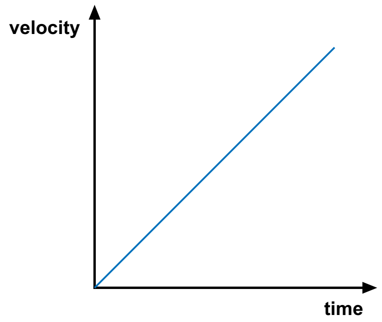
Find the time taken for a ball dropped from a height of 6.0 metres:

s (m)	u (m/s)	v (m/s)	a (m/s^2)	t (s)
6.0	0.0 (as dropped)	unknown	9.81	?

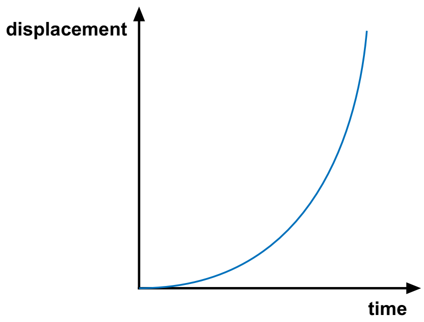
We don't know the final velocity so we must use equation 3 (because there is no v in this equation).

DID YOU KNOW?

Due to the shape of the Earth and its uneven density, the value of the acceleration due to gravity varies slightly. At the Equator it is 9.78 m/s^2 whereas at the North Pole it is 9.83 m/s^2. 9.81 m/s^2 is usually referred to as Standard Gravity.



a



b

Figure 1.5 *Graphs showing the vertical velocity and displacement as the ball falls.*

DID YOU KNOW?

If we ignore air resistance, the mass of an object does not affect the rate at which it accelerates. Galileo Galilei was the first to realise this back in the 17th century. At the end of the last Apollo 15 moon walk, Commander David Scott tested this theory. He dropped a hammer and a feather at the same time. As the surface of the moon is a vacuum, there was no air resistance and the feather fell at the same rate as the hammer.

$s = ut + \frac{1}{2}at^2$

$ut = 0$ as the ball was dropped so the equation becomes $s = \frac{1}{2}at^2$

This can be rearranged to $t = \sqrt{\frac{2s}{a}}$

$t = \sqrt{\frac{2 \times 6.0 \text{ m}}{9.81 \text{ m/s}^2}}$

$t = 1.1$ s (to two significant figures)

Activity 1.2: Dropping a ball

Drop a ball from several different heights and time how long it takes to hit the ground. Record your data carefully and take repeat measurements for each height.

Using the equation, calculate the time it actually takes to hit the ground. Compare the actual times with your readings and comment on your findings.

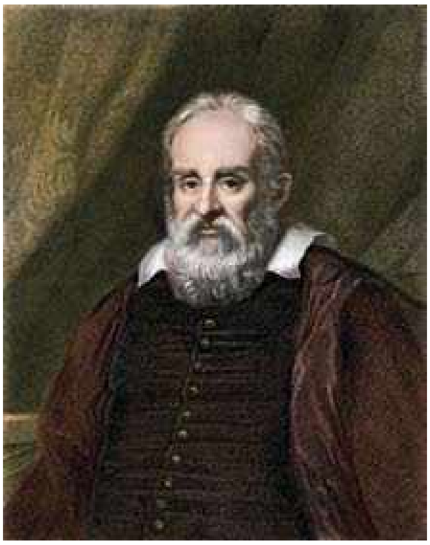


Figure 1.6 *Galileo Galilei.*

We can also work out the final vertical velocity.

Worked example 1.2

Find the final v ertical velocity for a ball dropped from 6 m. Looking back at the table we now have:

s (m)	u (m/s)	v (m/s)	a (m/s^2)	t (s)
6.0	0.0 (as dropped)	unknown	9.81	1.1

We could use equations 1, 2, 4, or 5 to determine v. However, equation 4 does not require you to calculate time, so this is preferable.

$v^2 = u^2 + 2as$

$v = \sqrt{(u^2 + 2as)}$

$v = \sqrt{(0^2 + 2 \times 9.81 \text{ m/s}^2 \times 6.0 \text{ m})}$

$v = 11$ m/s (to two significant figures)

KEY WORDS

parabola *the curved path a projectile takes through the air*

How does this affect the motion of projectiles?

When a projectile moves through the air, it follows a path caused by the combination of its horizontal and vertical velocities. This path is curved; it forms a special type of curve called a **parabola**.

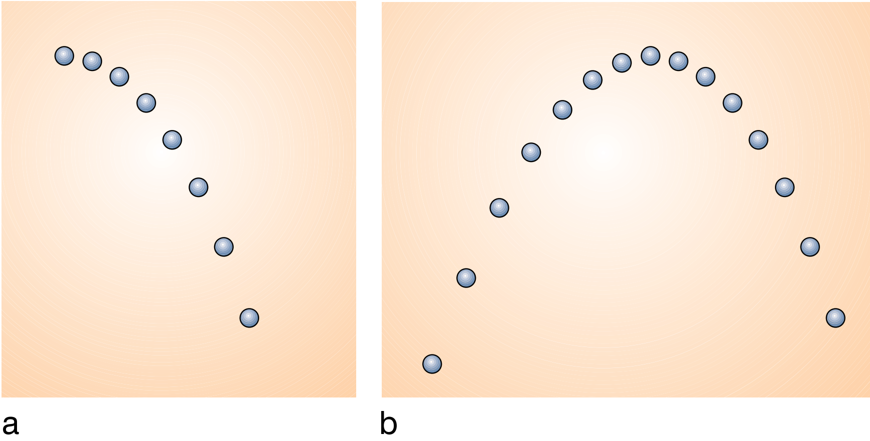


Figure 1.7 *Diagrams of paths of balls.*

Figure 1.9 shows a real example of this type of motion. Here one child throws a ball to another. A camera took photos at regular time intervals. If you look carefully, you can see the ball's vertical velocity increases as it falls and decreases as it rises.

Notice:

- Horizontally, the ball moves at a steady speed. The images of each ball are equally spaced horizontally.
- Vertically, the ball accelerates downwards due to gravity. This means the images become further and further apart.

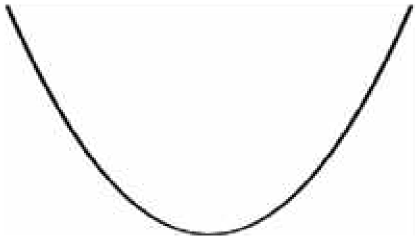


Figure 1.8 *A parabola.*

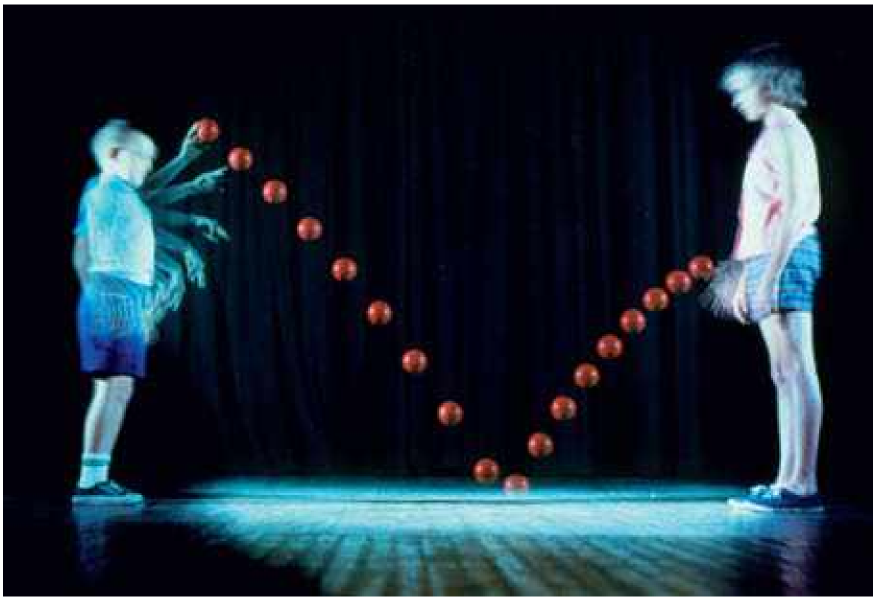


Figure 1.9 *A multi-flash image showing the motion of a projectile (a ball).*

Activity 1.3: The path of a ball

Next time you are outside, get two friends to throw a ball to each other. Stand at the side (some distance away) and watch the path of the ball carefully. Try asking your friends to throw the ball at different angles. What do you notice?

Horizontal and vertical velocities

This special shape is caused by the relationship between the horizontal and vertical velocities.

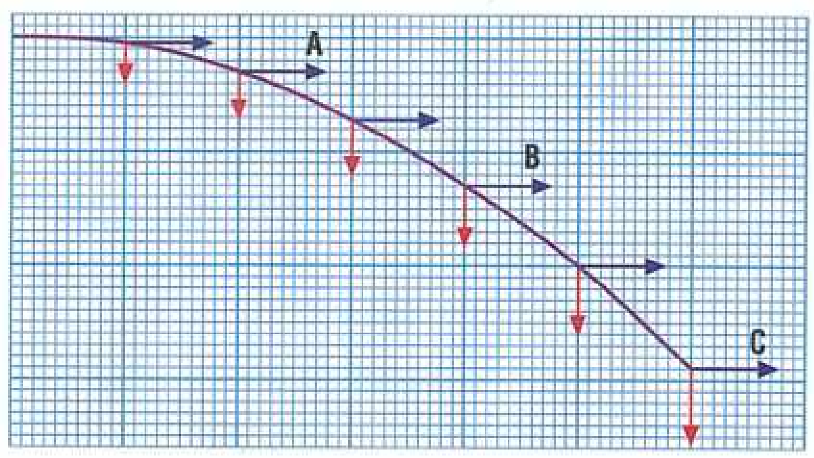


__Figure 1.10__ Horizontal and vertical velocities.

If you look carefully at Figure 1.10 you can see the horizontal velocity remains the same but the vertical velocity increases. This causes the ball to follow a parabolic path.

The same is true for a projectile thrown at an angle.

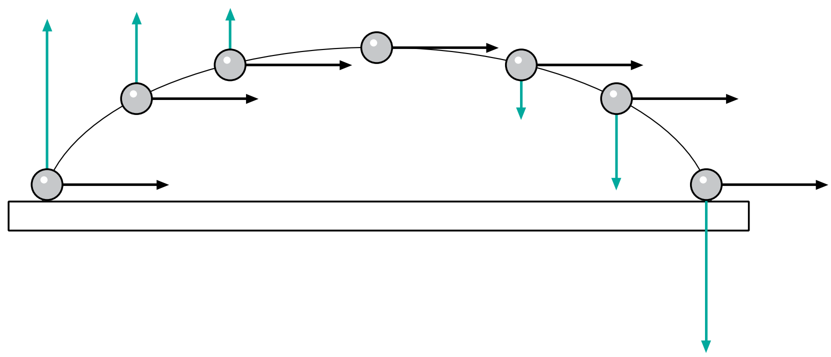


__Figure 1.11__ Horizontal and vertical velocities for a projectile thrown at an angle.

Figure 1.11 shows the velocities for a ball thrown at an angle. Again the horizontal velocity remains constant. The vertical velocity first decreases (on the way up), reaches zero (at the top of the flight), and then increases on the way down. The vertical velocity at any point is given by the equation $v = u + at$.

You need to think carefully about the directions of the vertical velocities and the acceleration. If we use a as 9.81 m/s^2, then the initial vertical velocity must be a negative number as it is in the opposite direction. We have effectively decided that downwards is the positive direction.

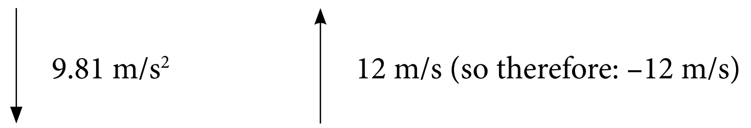


__Figure 1.12__ The importance of direction when dealing with vectors.

This works the other way around, too. It does not matter which way is positive and which way is negative, but you must not mix them up!

Worked example 1.3

An arrow is fired vertically with an initial velocity of 35 m/s. Find its velocity after:

a) 3 s

b) 6 s

Using the table layout seen earlier we get:

s (m)	u (m/s)	v (m/s)	a (m/s^2)	t (s)
Unknown	–35	?	9.81	a) 3 b) 6

Notice we have entered –35 m/s for the initial velocity. We are therefore setting the downwards direction as positive.

a)

$v = u + at$

$v = -35 \text{ m/s} + (9.81 \text{m/s}^2 \times 3 \text{ s})$

$v = -5.6$ m/s (to 2 significant figures)

Notice the velocity is still negative as it is still travelling upwards.

b)

$v = u + at$

$v = -35 \text{ m/s} + (9.81 \text{m/s}^2 \times 6 \text{ s})$

$v = 24$ m/s (to 2 significant figures)

Notice the velocity is now positive. This must mean the arrow has changed direction and is heading back down.

Activity 1.4: Heights of arrows

Calculate the height of the arrow in each case for the worked example.

Horizontal projection

Projectiles may be initially travelling horizontally. This might include a ball kicked off a wall, a bullet fired from a horizontal gun, or a parcel dropped from the underside of an aircraft flying horizontally.

The object will follow the path shown in Figure 1.13. It is interesting to note that the time it takes to hit the floor only depends on the original height of the object.

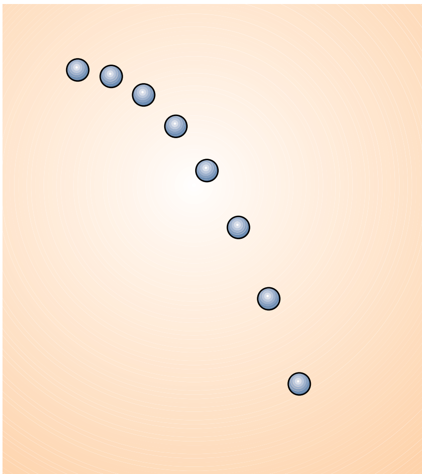


Figure 1.13 Horizontal projection.

The flight time is given by the equation $s = ut + \frac{1}{2}at^2$. Looking at this vertically, $ut = 0$ as the ball initially has no vertical velocity. So the equation becomes $s = \frac{1}{2}at^2$ and the time it takes to hit the floor is given by:

$$t = \sqrt{\frac{2s}{a}}$$

The acceleration is constant (9.81 m/s^2) and so the only factor affecting the time to hit the floor is the drop height.

From this we can draw some counterintuitive conclusions: imagine someone holding a rifle in their right hand and a rifle bullet in their left hand. The rifle is perfectly horizontal. The rifle is then fired and the bullet is dropped at exactly the same time. Which bullet will hit the ground first?

It would be tempting to say the bullet from the rifle takes longer to hit the ground than the one dropped from the hand. However, this would be wrong! They both hit the ground at the same time; the only difference is the bullet from the rifle hits the ground several metres away whereas the dropped bullet hits the ground by the person's feet.

Figure 1.14 shows a time lapse of two balls: one dropped, the other fired horizontally. Both were released simultaneously. You can clearly see they stay at exactly the same height as they fall.

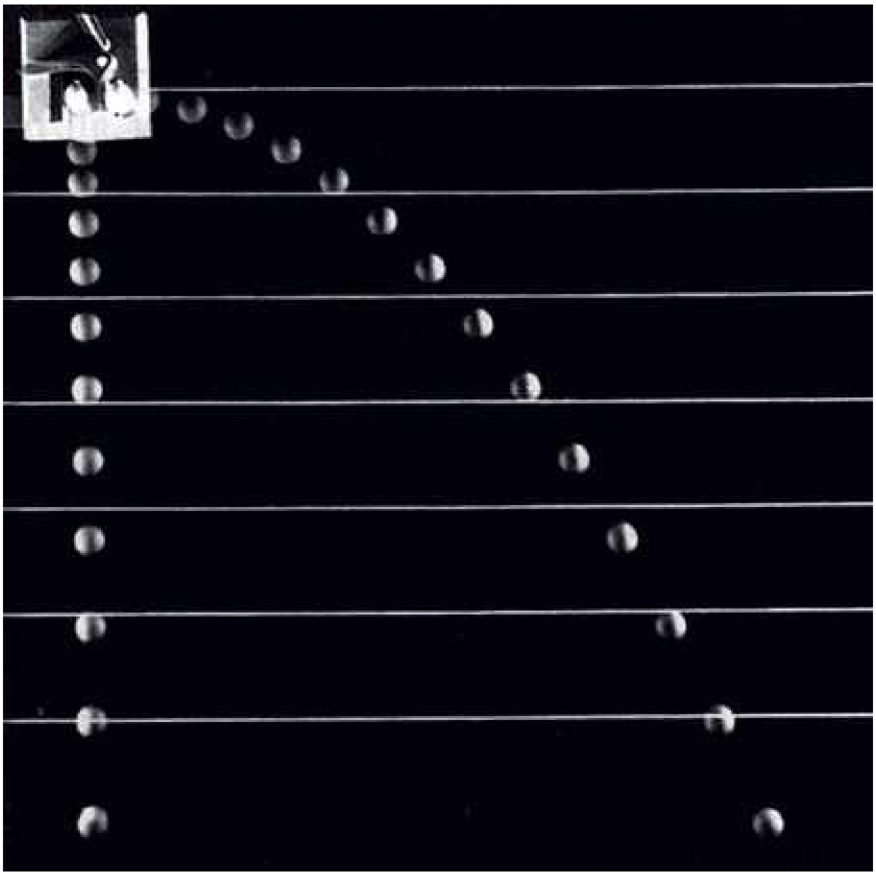


Figure 1.14 *Time lapse of two balls falling.*

Activity 1.5: Test with ruler and two coins

You can test this using a ruler and two coins. Carefully place the ruler on the edge of a desk. Place one coin on the end of the ruler overhanging the desk. Place the other next to the ruler as shown. The aim is to flick the ruler so that the first coin falls vertically where as the second coin gets pushed off the desk horizontally.

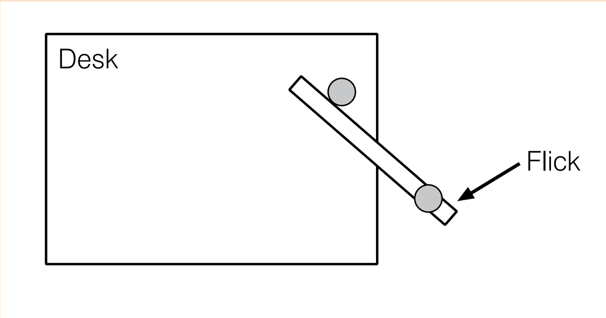


Figure 1.16 *Ruler and coin experiment.*

You have to flick it quite hard.

When you do so, listen for the clink as the coins hit the floor. You will find the two clinks come at once; both coins hit the floor at the same time.

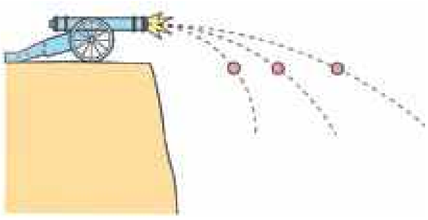


Figure 1.15 *Cannonballs fired horizontally with different velocities.*

Activity 1.6: Rolling a ball down a track

A more complex experiment involves rolling a ball or marble down a track. As you vary the release height, you vary the horizontal velocity of the marble. You could time the time it takes to hit the ground for different release heights.

You must be careful to start timing when the ball leaves the desk.

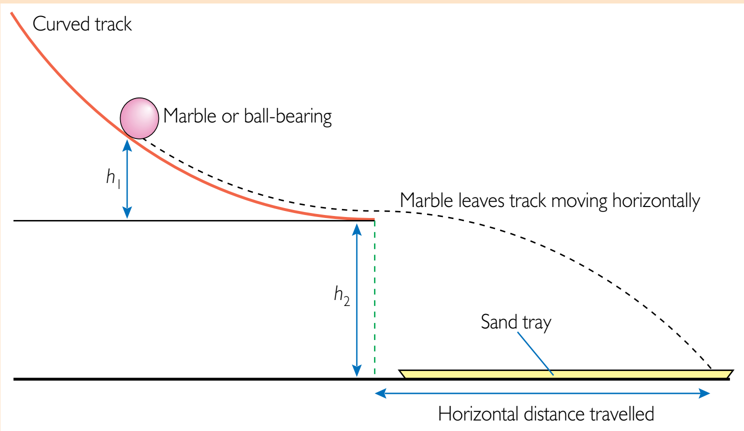


Figure 1.17 *A marble rolling down a track.*

Worked example 1.4

A football is kicked off a wall at an initial horizontal velocity of 12 m/s. The wall is 2.1 m high. Find the time taken for the ball to hit the floor and the range of the ball.

s (m)	**u** (m/s)	**v** (m/s)	**a** (m/s²)	**t** (s)
2.1	12	?	9.81	?

$t = \sqrt{\dfrac{2s}{a}}$

$t = \sqrt{\dfrac{2 \times 2.1\ \text{m}}{9.81\ \text{m/s}^2}}$

$t = 0.65$ s

horizontal displacement = horizontal velocity × flight time

horizontal displacement = 12 m/s × 0.65 s

horizontal displacement = 7.8 m

As discussed earlier, the range of any projectile is given by:

horizontal displacement = horizontal velocity (v_h) × flight time

The flight time is given by: $t = \sqrt{\dfrac{2s}{a}}$

Combining these we get:

$$\text{horizontal displacement} = v_h \times \sqrt{\frac{2s}{a}}$$

$$\text{horizontal displacement} = v_h \times \sqrt{\frac{2\ \text{vertical height}}{\text{acceleration}}}$$

This only applies to projectiles initially travelling horizontally.

Activity 1.7: Trajectory of a ball

Use the equations above to complete the table below. Use this data and a piece of graph paper to carefully plot the trajectory of a ball thrown horizontally with a velocity of 2.0 m/s.

Time (s)	**Vertical displacement *y* (m)**	**Horizontal displacement *x* (m)**
1.0		
2.0		
3.0		
4.0		
5.0		

Activity 1.8: Range of a ball rolled down a track

Using the same equipment as for Activity 1.6, roll the marble down the track. This time, measure the range of the marble when released from several different heights up the ramp. Be sure to repeat your readings for each height.

Plot a graph of your findings and use your data to calculate the initial velocity in each case.

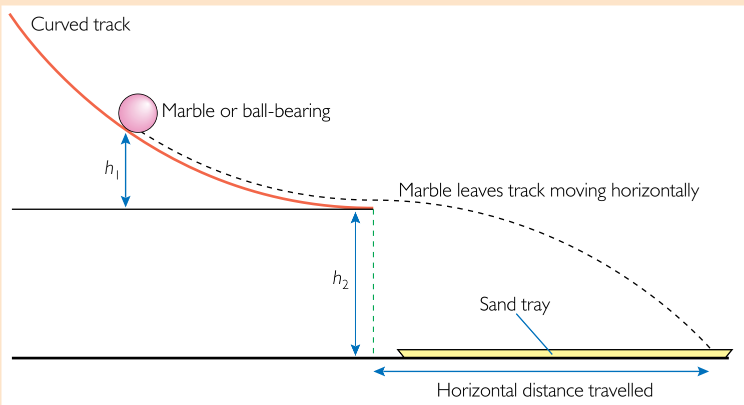


***Figure 1.18** A marble rolling down a track.*

Projectiles at angles

Projectiles may also be travelling at angles to the horizontal. In this case their initial vertical velocity is not zero. In order to find the initial horizontal and initial vertical velocity, the velocity must be resolved into horizontal and vertical components.

For example, a ball kicked with a velocity of 9.0 m/s at an angle of 50° to the horizontal:

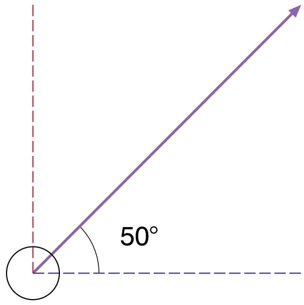


Figure 1.19 *Ball kicked at an angle.*

To find the components, we use trigonometry.

- initial vertical component = 9.0 m/s × sin 50°
- initial vertical component = 6.9 m/s
- initial horizontal component = 9.0 m/s × cos 50°
- initial horizontal component = 5.8 m/s (remember this will not change throughout the flight of the projectile).

If the angle of the projectile is above the horizontal, then this is referred to as an **angle of elevation** (Figure 1.20a). However, if the angle of the projectile is below the horizontal then this is referred to as an **angle of depression** (Figure 1.20b).

Using the initial vertical velocity, the initial horizontal velocity and the equations of uniform acceleration we can then determine the range, flight time and maximum height of the projectile.

KEY WORDS

angle of elevation *the angle of a projectile's trajectory above the horizontal*

angle of depression *the angle of a projectile's trajectory below the horizontal*

Maximum height

At the maximum height the vertical velocity of the projectile will be zero. Using the example in Figure 1.19 (9.0 m/s at 50°) we can use the equations of uniform acceleration as follows:

s (m)	u (m/s)	v (m/s)	a (m/s²)	t (s)
??	$u \sin \theta$ in this case: −6.9	At max height: 0	9.81	unknown

Again, take care to ensure you consider the directions of the velocities; in this case, downwards is positive.

To find the maximum height we use $v^2 = u^2 + 2as$ (as we don't know t):

$$v^2 = u^2 + 2as$$

$$s = \frac{v^2 - u^2}{2a}$$

$$s = \frac{(0^2 - (-6.9^2\ \text{m/s}^2)}{2 \times 9.81\ \text{m/s}^2}$$

$$= -2.4\ \text{m.}$$

The height is of course 2.4 m, the negative just indicates that it is in the opposite direction to the acceleration.

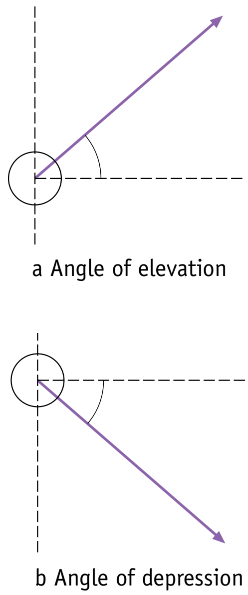


Figure 1.20 *Angle of elevation and angle of depression.*

We can then work out the time taken to reach maximum height using $v = u + at$.

$$v = u + at$$

$$t = \frac{v - \mathrm{u}}{a}$$

$$t = \frac{0 - -6.9\ \mathrm{m/s}}{9.81\ \mathrm{m/s^2}}$$

$$t = 0.70\ \mathrm{s}$$

Projectile range

If the projectile is launched from the ground, the total range will be given by:

- Total horizontal displacement = horizontal velocity × total flight time

The horizontal velocity remains the same throughout the flight and is given by $u \cos \theta$ (in the previous example, 5.8 m/s). All we need to find the range is total flight time.

This can be done in a number of ways; perhaps the simplest is to use $v = u + at$. The final vertical velocity will be the same magnitude as the initial vertical velocity, just in the opposite direction (look at Figure 1.11 to check).

Using the same example as before we get:

s (m)	u (m/s)	v (m/s)	a (m/s^2)	t (s)
??	$u \sin \theta$ in this case: −6.9	6.9	9.81	unknown

Again, take care to ensure you consider the directions of the velocities; in this case downwards is positive. Notice the final velocity is the same magnitude as the initial velocity, just the opposite direction (hence positive and not negative).

We can then work out the time taken to reach maximum height using $v = u + at$.

$$v = u + \mathrm{at}$$

$$t = \frac{v - u}{\mathrm{a}}$$

$$t = \frac{6.9\ \mathrm{m/s} - -6.9\mathrm{m/s}}{9.81\ \mathrm{m/s^2}}$$

$$t = 1.4\ \mathrm{s}$$

The total flight time is just double the time to reach maximum height!

So the range of the football is:

Total horizontal displacement = horizontal velocity × total flight time

Total horizontal displacement = 5.8 m/s × 1.4 s

Total horizontal displacement = 8.1 m

Alternatively, you can use algebra to combine the equation above, the horizontal and vertical components of the velocity, and the equations of uniform motion. This gives an expression for the range as:

$$\text{range} = \frac{u \cos\theta \times 2u \sin\theta}{a}$$

or

$$\text{range} = \frac{2u^2 \sin\theta \times \cos\theta}{a} = \frac{u^2 \sin 2\theta}{a}$$

This is usually called the **range equation**.

Worked example 1.5

Find the range of a projectile launched at an angle of 50° with an initial velocity of 30 m/s.

range (m)	u (m/s)	a (m/s²)	θ (°)	2θ (°)	sin 2θ
?	30	9.81	50°	100°	0.9848

Use $\text{range} = \frac{u^2 \sin 2\theta}{a}$

$$\text{range} = \frac{30 \times 30 \times 0.9848}{9.81}$$

$$= 90.35 \text{ m}$$

Flight time

You can also derive an equation for the total flight time for a projectile fired at an angle. This is just a version of $s = vt - \frac{1}{2}at^2$ (equation 5 in our list of equations for constant acceleration).

At the end of the flight the vertical displacement will be 0 m, as the object will be back on the ground. The final vertical velocity will be given by $u \sin\theta$, as the object ends up with the same vertical velocity as it started, just in the opposite direction. So:

$$s = vt - \tfrac{1}{2}at^2$$

$$0 = (u \sin\theta)\, t - \tfrac{1}{2}at^2$$

Activity 1.9

In a small group, discuss why

$$\frac{2u^2 \sin\theta \times \cos\theta}{a} = \frac{u^2 \sin 2\theta}{a}$$

KEY WORDS

range equation *an algebraic expression to calculate the range of a projectile*

Rearranging this we get:

$\frac{1}{2}\ a\ t^2 = (u \sin \theta)\ t$

$\frac{1}{2}\ a\ t = u \sin \theta$

$a\ t = 2\ u \sin \theta$

$$t = \frac{2u \sin \theta}{a}$$

Worked example 1.6

Find the flight time for a cannonball launched with a velocity of 60 m/s at an angle of 30°.

t (s)	u (m/s)	θ (°)	sin θ (°)	a (m/s²)
?	60	30	0.5	9.81

$$t = \frac{2u \sin \theta}{a}$$

$$t = \frac{2 \times 60 \sin 30}{9.81}$$

$t = 6.1$ s

Its range would be:

$$\text{Range} = \frac{2u^2 \sin \theta \times \cos \theta}{a}$$

$$\text{Range} = \frac{2 \times 60^2 \sin 30 \times \cos 30}{9.81}$$

Range = 320 m

Remember that all of these equations ignore air resistance. The equations become exceptionally complex when this is factored in, especially since the air resistance changes as the velocity changes.

Maximum range

The angle of a projectile affects its maximum range.

Figure 1.21 shows different paths of a projectile fired at different angles. The maximum range is achieved when the angle is 45°. This can be explained be referring back to the range equation:

$$\text{Range} = \frac{2u^2 \sin \theta \times \cos \theta}{a}$$

The maximum value of sin $\theta \times$ cos θ is 0.5 and this happens when θ is 45°. This gives us:

$$\text{maximum range} = \frac{2u^2\ 0.5}{a}$$

$$\text{maximum range} = \frac{u^2}{a}$$

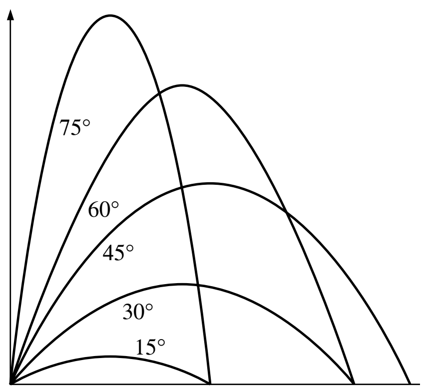


Figure 1.21 An illustration of the effect of angle on range.

To determine the range of projectiles launched above the ground you need to use algebra to derive a new equation. The range is given by:

$$\text{Range} = \frac{(u\cos\theta) \times (u\sin\theta + \sqrt{((u\sin\theta)^2 + 2ah)}}{a}$$

where h is the height above the ground. Even this equation ignores air resistance!

Activity 1.10: The range of a cannonball

Use the range equation to determine the range of a cannonball fired with a velocity of 50 m/s when fired at a series of different angles. Use 15°, 30°, 45°, 60° and 75°. Plot a graph of range against angle and comment on your findings.

Worked example 1.7

Use this equation to determine the range of the ball in Figure 1.22.

u (m/s)	θ (°)	cos θ (°)	sin θ	a (m/s²)	h (m)	range (m)
36	39	0.7771	0.6293	9.81	1.6	?

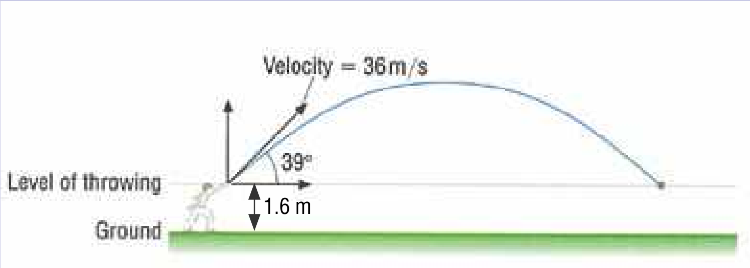


***Figure 1.22** Throwing a soccer ball.*

$$\text{Use range} = \frac{(u\cos\theta) \times (u\sin\theta) + \sqrt{((u\sin\theta)^2 + 2ah)}}{a}$$

$$= \frac{(36 \times 0.7771) \times (36 \times 0.6293) + \sqrt{((36 \times 0.6293)^2 + 2 \times 9.81 \times 1.6)}}{9.81}$$

$$= \frac{27.9756 \times (22.6548) + \sqrt{((22.6548)^2 + 31.392)}}{9.81}$$

$$= \frac{27.9756 \times (22.6548) + \sqrt{(513.24) + 31.392)}}{9.81}$$

$$= \frac{27.9756 \times (22.6548) + \sqrt{(544.632))}}{9.81}$$

$$= \frac{27.9756 \times (22.6548) + \sqrt{(544.632))}}{9.81}$$

$$= \frac{27.9756 \times (22.6548) + (23.3374)}{9.81}$$

$$= \frac{27.9756 \times 45.9922}{9.81} = \frac{1286.66}{9.81} = 131.16 \text{ m}$$

KEY WORDS

centre of mass *the point at which all the mass of an object may be considered to be concentrated*

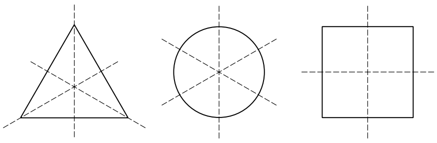


Figure 1.23 *Examples of the centre of mass for difference uniform shapes.*

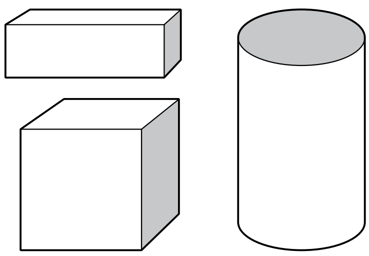


Figure 1.24 *The centre of mass for different uniform 3D objects is in the centre of the object.*

Centre of mass

All objects have a **centre of mass**. This is the point at which all the mass of the object may be considered to be concentrated.

For uniform objects the centre of mass will be at the intersection of all the lines of symmetry; in essence the centre of the object.

This is also true of 3D objects.

Another way to define the centre of mass is: the point through which a single force on a body has no turning effect.

Activity 1.11: Centre of mass of a ruler

Take a ruler and balance it on your finger. When balanced, the centre of mass must be above your finger. Take several other objects and balance them on your finger to find the approximate position of the centre of mass. Try it with a few more irregularly shaped objects.

This idea of balancing an object so there is no net turning effect leads onto the centre of mass theorem. This is a mathematical treatment of the distribution of mass, which is beyond the scope of this course.

The centre of mass theorem

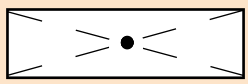
The centre of mass theorem simply states:

- when a force is applied to an object, the object acts as though its mass were a point mass at its centre of mass.

This means that the motion of the centre of mass of a system is identical to the motion of a single particle with the same mass as the system if such a particle were acted on by the same external forces.

Activity 1.12: Motion of the centre of mass

Work in a small group. You need a 2 m plank with four low friction wheels (for example, roller bearings).



- Mark the middle of the plank as shown in the diagram (this is the centre of mass), on the plank and on the floor beneath.
- One person should stand at the middle of the plank and walk four steps.
- Mark the position where the student finishes on the plank and on the floor.
- Record the distance that the centre of mass has moved.
- Use your knowledge of motion to analyse the motion of the centre of mass.

The centre of mass where two or more celestial bodies orbit each other is known as the barycentre. This is the point between the bodies where they balance each other. The Moon does not orbit the exact centre of the Earth, but their masses balance at a point approximately 1710 km below the surface of the Earth on a line between the Earth and the Moon.

Experimental determination of centre of mass

It is quite difficult to accurately determine the centre of mass for a 3D object. Special machines called planimeters are used. However, it is quite simple to determine the mass of a 2D object.

Activity 1.13: Centre of mass of an object

1. Take a piece of thick card and cut it into any shape you like (this example involves using a piece shaped like a jigsaw piece).
2. Make a series of small holes around the edge of the shape.
3. Hang it from one of these holes so the object is free to rotate.
4. Construct a simple plumb line using some wire (or string) and a mass.
5. Hang this from the support so it hangs vertically down.
6. Using a sharp pencil, draw a line to show the position of the plumb line.
7. Repeat this procedure for all the holes you have made, making a series of straight lines on your shape.
8. The lines should all cross; this is the centre of mass of the object. (You can test this by balancing the shape on the sharp pencil.)

The centre of mass of a system does not have to be an object. Take, for example, a cup. The centre of mass will be inside the cup even though there is nothing there but air. A more complex example might be two binary stars; these orbit the centre of mass between the two stars.

Uniform circular motion

Another example of 2D motion is **uniform circular motion**. This does not just refer to objects spinning around in circles, but also to objects following a curved path that is the shape of part of a circle, such as a car going around a bend of constant radius.

- Uniform circular motion specifically refers to following a curved path of constant radius at a steady speed.

In this case, despite the speed remaining constant, the velocity is constantly changing. This is because velocity is a vector quantity, so as the object moves around the circle its direction is changing, and therefore its velocity must be changing. This can be seen in Figure 1.27, where the velocity has changed between points A and B.

DID YOU KNOW?

The terms centre of gravity and centre of mass are often confused. There is a slight technical difference, but this is only apparent if the object is in a non-uniform gravitational field.

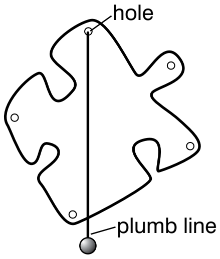


***Figure 1.25** Determining the centre of mass of a 2D object.*

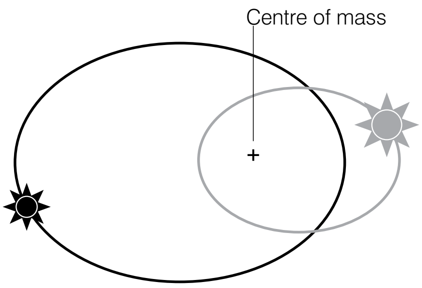


***Figure 1.26** Centre of mass of two binary stars.*

KEY WORD

uniform circular motion *the motion of a body following a curved path of constant radius at constant speed*

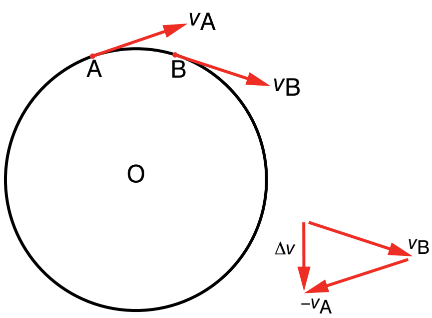


***Figure 1.27** The changing velocity for an object moving in a circular path.*

KEY WORDS

centripetal force *the net force needed to make a body follow a circular path*

As the velocity is changing, the object must be accelerating (acceleration is defined as the rate of change of velocity). According to Newton's second law, acceleration requires a net force and the direction of the acceleration is the same direction as the net force.

Projectile motion is not an example of uniform circular motion. As the path is a parabola, the radius is constantly changing.

Figure 1.28 shows the direction of the net force required to keep an object following a circular path. This force is referred to as the **centripetal force** and it always acts towards the centre of the circle.

What factors affect the size of the centripetal force?

The mass of the object, its velocity, and the radius of the curved path followed by the object affect the centripetal force.

Activity 1.14: A simple pendulum

Make a simple pendulum using a mass and a piece of string. *Carefully* swing this around your head in a horizontal circle, trying to keep the speed constant.

The force required to make the mass move and follow a circular path will come from your hand. You will be able to feel the force required. Experiment by changing the mass, radius and velocity of the object. How does this affect the force required?

To calculate the centripetal force we use the equation below:

$$F = \frac{mv^2}{r}$$

where F = centripetal force, m = mass of the object,
v = velocity of the object, and r = radius of the curved path.

Worked example 1.8

The mass of the Earth is 6.0×10^{24} kg. It travels at a steady speed around the Sun at 30 000 m/s at a radius of 1.5×10^{9} m. Find the force required to keep it in orbit.

F **(N)**	***m*** **(kg)**	***v*** **(m/s)**	***r*** **(m)**
?	6×10^{24}	30 000	1.5×10^{9}

$$F = \frac{mv^2}{r}$$

$$F = \frac{6.0 \times 10^{24}\ \text{kg} \times 30\ 000\ \text{m/s}^2}{1.5 \times 10\ \text{m}}$$

$$F = 3.6 \times 10^{24}\ \text{N}$$

Circular motion examples

There are plenty of examples of circular motion. In each case the centripetal force acts towards the centre of the circular path followed by the object.

Table 1.1 *Examples of centripetal forces.*

Context	Centripetal force	Direction of the force
The Earth orbiting the Sun	Gravitational attraction	Towards the centre of the circular path: towards the centre of the Sun
An electron orbiting an atom	Electrostatic attraction	Towards the centre of the circular path: towards the centre of the nucleus
A ball on a string being whirled around	Tension in the string	Towards the centre of the circular path: towards the centre of the circle
A car going around a bend	Friction	Towards the centre of the circular path: towards the centre of the bend
A bus going over a hump-backed bridge	A component of weight	Towards the centre of the circular path: towards the centre of the bridge

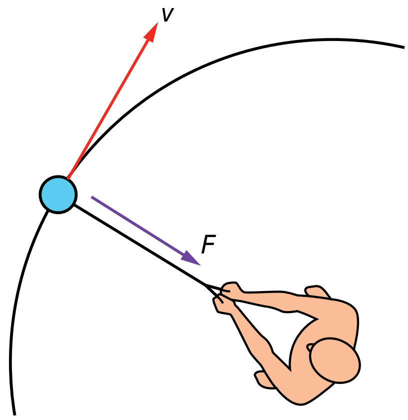


Figure 1.28 *Forces acting on a hammer thrower.*

Worked example 1.9

A hammer thrower swings a hammer around his head at a steady speed of 2.0 m/s at a radius of 1.2 m. He exerts a constant force of 120 N, Calculate the mass of the hammer.

F (N)	*m* (kg)	*v* (m/s)	*r* (m)
120	?	2.0	1.2

$$F = \frac{mv^2}{r}$$

$$Fr = mv^2$$

$$m = \frac{Fr}{v^2}$$

$$m = \frac{120\ \text{N} \times 1.2\ \text{m}}{2\ \text{m/s}^2}$$

$$m = 36\ \text{kg}$$

What if the centripetal forces are not large enough?

There is a maximum centripetal force that can be provided. For example, if the tension gets too high, the string will snap. Likewise, if the friction is not high enough the car may skid or slide.

Imagine the maximum frictional force between the road and the tyres of a certain car is 6500 N. The mass of the car is 1200 kg and the bend has a radius of 85 m. Determine the maximum speed at which the car can take the bend.

$$F = \frac{mv^2}{r}$$

$$v^2 = \frac{Fr}{m}$$

$$v = \sqrt{\frac{Fr}{m}}$$

$$v = \sqrt{\frac{6500\ \text{N} \times 85\ \text{m}}{1200\ \text{kg}}}$$

$$v = 21\ \text{m/s}$$

Discussion activity

Look carefully at the centripetal force equation. What effect does changing each of the variables have on the force required? For example, what would happen to the required force if the mass doubled? What would happen if the velocity doubled?

If the car travels faster than this, friction will not be large enough to provide the required centripetal force. The car will then follow a path of increased radius, and skid towards the edge of the road.

Summary

In this section you have learnt:

- A projectile is any object moving through the air without an engine or other motive force; examples include tennis balls and rifle bullets.
- Projectile motion and uniform circular motion are examples of 2D motion.
- Gravity causes projectiles to follow a parabolic path; this is because the horizontal velocity remains the same but the vertical velocity increases.
- The angle of a projectile affects its flight path and therefore its range and the time in the air.
- The equations for uniform acceleration can be applied to projectile motion.
- Centre of mass is the point at which all the mass of the object may be considered to be concentrated.
- Uniform circular motion is when an object travels at a steady speed around a circular path.

Review questions

1. This question is about a simple model of the physics of the long jump. Figure 1.29 shows a long-jumper at three different stages, A, B and C, during the jump. The horizontal and vertical components of velocity at each position are shown.

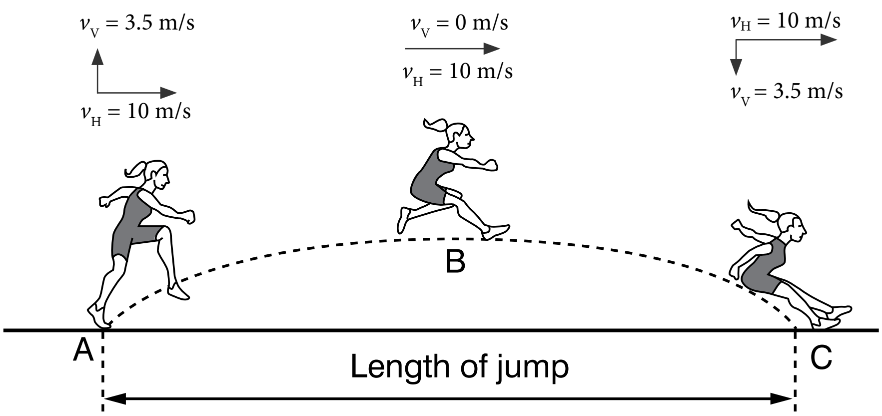


Figure 1.29

a) i) In the model, the horizontal component of velocity v_H is constant at 10 m/s throughout the jump. State the assumption that has been made in the model.

ii) Without calculation, explain why the vertical component of velocity v_H changes from 3.5 m/s at A to 0 m/s at B.

b) i) By considering only the vertical motion, show that it takes about 0.4 s for the jumper to reach maximum height at B after taking off from A. Take $g = 9.8$ m/s^2.

ii) Calculate the length in metres of the jump.

c) Long-jumpers can use this model to help them to improve their performance. Explain why the length of the jump can be increased by:

i) increasing the horizontal component of velocity v_H, keeping v_V the same.

ii) increasing the vertical component of velocity v_V, keeping v_H the same.

2. Define the term centre of mass.

3. Describe in as much detail as you can the motion of a projectile fired at an angle (include descriptions of the vertical and horizontal components of velocity).

4. A football is kicked at a velocity of 15 m/s at an angle of 25° to the horizontal. Calculate:

a) the total flight time

b) the maximum height

c) the range of the ball.

5. A car accidently rolls off a cliff. As it leaves the cliff it has a horizontal velocity of 13 m/s, it hits the ground 60 m from the shoreline. Calculate the height of the cliff.

6. A car of mass 1200 kg travels at a steady speed of 22 m/s around a bend of radius 50 m. Find the centripetal force required.

7. Find out what is meant by the centre of stability of a ship. Why is it important that the centre of mass of a ship is below the centre of stability?

8. A tennis ball machine fires a ball vertically upwards at time $t = 0$ at 19.6 m/s. Assume that air resistance is negligible and take $g = 9.8$ m/s^2.

a) Calculate the displacement and velocity of the ball at $t = 1.0$ s, 2.0 s, 3.0 s and 4.0 s.

b) Use your answers to part a) to draw

i) a graph of displacement against time for the ball

ii) a graph of velocity against time for the ball.

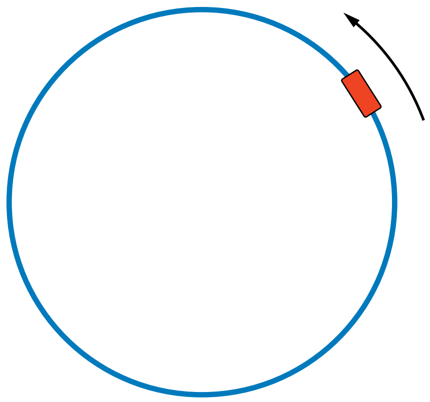


Figure 1.30 Example of circular motion.

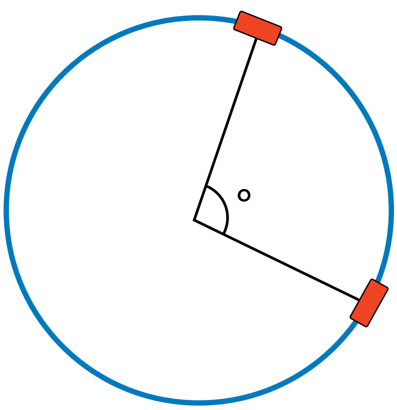


Figure 1.31 Angular displacement.

1.2 Rotational kinematics

By the end of this section you should be able to:

- Define the terms angular and tangential displacement, and angular and tangential velocity.
- Express angles in terms of revolutions, radians and degrees.
- Define the term angular acceleration, and list its key characteristics.
- Identify the SI unit of angular velocity and angular acceleration.
- Explain the relationships between angular displacement, tangential displacement, angular velocity, tangential velocity, angular acceleration and tangential velocity.
- Demonstrate how to use the equations of constant angular acceleration and compare them with equations of constant acceleration.

Rotational motion

Whenever objects travel in curved paths their motion can be considered to be rotational. A simple example might be a ball on a piece of string being swung around in a circular path. A more complex example might include the orbit of the planets around the Sun.

Up until this point, whenever we have used the terms displacement, velocity and acceleration this has always applied to a straight line. We now need to consider the motion of objects following curved paths. Let's take an example of an object travelling in a perfect circle.

As discussed earlier, this object is accelerating (it is changing direction, therefore changing velocity, therefore accelerating). However, what if the object was also getting faster? We need to be able to distinguish between its displacement, velocity and acceleration, and whether it is angular or tangential.

Displacement

In one complete revolution an object will travel a distance equal to $2\pi r$ (where r is the radius of the circle). Its tangential displacement will be zero as it is back where it started. If we consider part of this motion, between two points, we can see that the object has an angular displacement equal to θ. This is just the angle the object has subtended.

When describing angular displacement, there are several different units of angular measurement we could use. Perhaps the simplest would be revolutions. Half a circle would be 0.5 revolutions (or 0.5 rev), two complete loops would be 2 revolutions. However, this does not have much scientific merit. A more common unit is the degree.

DID YOU KNOW?

Having 360° in a circle is convenient because it divides easily into a whole a number. It divides by 1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 18, 20, 24, 30, 36, 40, 45, 60, 72, 90, 120, 180, and 360!

Discussion activity

Why are there 360° in a circle? Discuss some possible reasons with a partner.

One revolution is 360°, half a circle would be 180°, and so on.

For scientific and mathematical calculations the **radian** is used. This has a clear mathematical basis and offers a number of advantages (especially when dealing with high level trigonometry). One radian is the angle subtended at the centre of a circle by an arc that is equal in length to the radius of the circle.

This definition means that there must be 2π radians in a circle. This is because the circumference of a circle is given by $2\pi r$ and so the radius fits around a complete circle 2π times. This could be written as 6.28 rad, but it is usually just written as 2π rad. Half a circle would be π rad (or 3.14 rad).

One radian is equal to $\frac{180}{\pi}$ (or $\frac{360}{2\pi}$) degrees, or about 57.3°.

Table 1.2 Examples of angular measurements

Revolutions	Degrees	Radians
$\frac{1}{8}$	45	$\frac{\pi}{4}$
$\frac{1}{4}$	90	$\frac{\pi}{2}$
$\frac{1}{2}$	180	π
$\frac{3}{4}$	270	$\frac{3\pi}{2}$
1	360	2π
2	720	4π

Velocity

Angular velocity is defined as the rate of change of the angle subtended. This is very similar to the definition of linear velocity (rate of change of displacement). The faster an object rotates, the greater the angle covered per unit of time. This leads us to the relationship:

$$\textit{average angular velocity} = \frac{\textit{angle covered}}{\textit{time taken}}$$

In symbols:

$$\omega = \frac{\theta}{t}$$

ω is the Greek letter omega. As θ is measured in radians and t in seconds, the units of angular velocity are therefore rad/s.

In one complete revolution θ is equal to 2π radians and t becomes T (time for one complete cycle).

So for one complete cycle:

$$\omega = \frac{\theta}{t} = \frac{2\pi}{T}$$

This allows us to work out angular velocities if we know the time taken for one cycle.

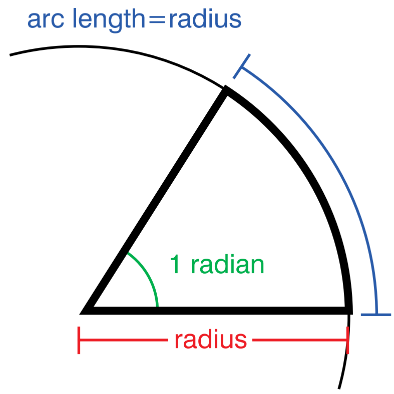


***Figure 1.32** The radian.*

Activity 1.15: Calculating angular displacement

Thinking about an analogue clock, calculate the angular displacement in revolutions, radians and degrees for the following:

- A second hand after three minutes
- An hour hand after 20 minutes
- The minute hand as it moves from 9.15 to 12.45.

KEY WORDS

radian *the angle between two radii of a circle that cut off, on the circumference, an arc equal in length to the radius*

angular velocity *specifies the angular speed of an object and the axis about which the object is rotating*

Activity 1.16: Calulating angular velocity

Other units of angular velocity could be degrees per second or even rpm (revolutions per minute). What would an angular velocity of 3π rad/s be in degrees per second and rpm?

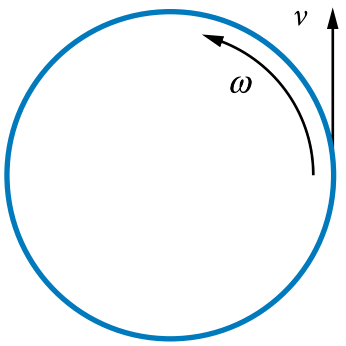


Figure 1.33 *Angular and tangential velocity.*

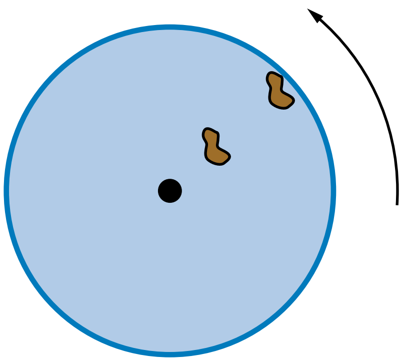


Figure 1.34 *Two stones on a spinning disk.*

KEY WORDS

acceleration *the rate of change of velocity as a function of time*

Worked example 1.10

The Earth takes approximately 365 days to orbit the Sun. Calculate its average angular velocity.

θ	*T*	*t*
2π	?	?

First we must find the time taken for one orbit in seconds

365 days = $365 \times 24 \times 60 \times 60 = 3.2 \times 10^7$ s

$$\omega = \frac{\theta}{t} = \frac{2\pi}{T}$$

$$= \frac{2\pi}{3.2} \times 10^7$$

$$= 2.0 \times 10^{-7} \text{ rad/s}$$

As the object follows a circular path it also has a tangential velocity at any given point (see Figure 1.33).

Angular and tangential velocity may be linked using the equation below:

tangential velocity = *radius of path* × *angular velocity*

$v = r\omega$

Using the Earth as an example, its tangential velocity at any one point may be calculated:

$v = r\omega$

$v = 1.5 \times 10^{11} \times 2.0 \times 10^{-7}$

$v = 30\,000$ m/s

Thinking about this equation further leads us to the conclusion that for any given angular velocity the tangential velocity is proportional to the radius.

Imagine two stones on a spinning disk (see Figure 1.34).

Both stones have the same angular velocity as they both cover the same angle in the same time. However, the stone nearer the edge has a higher speed (as it covers a greater distance in the same time) and so a greater tangential velocity.

Another way to think about uniform circular motion is to define it as any motion where both the speed and the angular velocity are constant.

Acceleration

As discussed in the previous section, any object travelling in a circular path is accelerating. This acceleration is called centripetal **acceleration** and it acts towards the centre of the circle.

Centripetal acceleration is given by the equation:

$$\text{Centripetal acceleration} = \frac{\text{tangential velocity}^2}{\text{radius of curvature}}$$

$$a = \frac{v^2}{r}$$

Since $v = r\omega$ this equation could be written as:

$$a = \frac{r^2\omega^2}{r}$$

$$a = r\omega^2$$

Again using the Earth as an example, its centripetal acceleration may be found using the following technique:

$$a = r\omega^2$$

$$a = 1.5 \times 10^{11} \times (2.0 \times 10^{-7})^2$$

$$a = 6.0 \times 10^{-3}\ \text{m/s}^2$$

Activity 1.17: Calculations using planets

Using the data in the table below, calculate the angular velocity, tangential velocity and centripetal acceleration for each of the planets listed.

Planet	Orbital period (days)	Average distance from Sun (m)
Mercury	88	5.8×10^{10}
Venus	225	1.1×10^{11}
Mars	686	2.3×10^{11}
Jupiter	4330	7.8×10^{11}
Neptune	60 000	4.5×10^{12}

If the mass of an object remains constant, then applying Newton's second law of motion gives us:

$F = ma$.

For centripetal acceleration and so centripetal force this equation becomes $F = \frac{mv^2}{r}$.

In the examples we have looked at so far, the object moving in a circular path has been travelling at constant speed. What if the object is also getting faster as it rotates? An example might be a car getting faster as it goes around a bend. In this case its tangential velocity is increasing, and as a result it has a tangential acceleration. The magnitude of this acceleration is given by:

$$\textit{tangential acceleration} = \frac{\textit{change in tangential velocity}}{\textit{time taken}}$$

$$a_T = \frac{\Delta v}{t}$$

In most cases an increase in tangential velocity will lead to an increase of both tangential and centripetal acceleration.

If the radius remains constant, an increase in tangential velocity will also cause an increase in angular velocity. Changes in angular velocity are as common as changes in linear velocity. Just think about a CD spinning up or the wheels of a car as it accelerates.

DID YOU KNOW?

Tangential and angular acceleration are related in the same way as tangential velocity and angular velocity. Instead of $v = r\omega$ we get $a_T = r\alpha$.

Discussion activity

Under what circumstances is it possible to have a constant centripetal acceleration but also be accelerating tangentially?

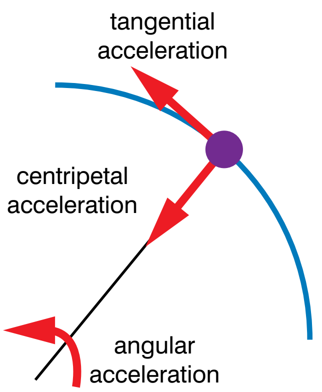


Figure 1.35 *Different accelerations in rotational motion.*

Angular acceleration is defined in a very similar way to linear acceleration:

$$\textit{angular acceleration} = \frac{\textit{change in angular velocity}}{\textit{time taken}}$$

$$\alpha = \frac{\Delta\omega}{t}$$

α is the Greek letter alpha. As ω is measured in rad/s and t in seconds, the units of angular acceleration are therefore rad/s^2.

This means that for an object travelling with rotational motion there may be three different types of acceleration.

Worked example 1.11

Find the final angular velocity when the initial angular velocity is 2 rad/s, the angular acceleration is 3 rad/s^2 and the time of motion is 4 s.

θ (rad/s)	ω_0 (rad/s)	α (rad/s^2)	t (s)
?	2	3	4

Use $\omega = \omega_0 + \alpha t$

Substitute in given values to get:

$\omega = 2 + 3 \times 4$

$= 14$ rad/s

The equations of constant angular acceleration

You saw on page 5 that there are equations that can be used when there is constant linear acceleration. There is an equivalent set of equations for constant angular acceleration.

Here are the two sets of equations side by side for comparison.

Constant linear acceleration	Constant angular acceleration
$v = u + at$	$\omega = \omega_0 + \alpha t$
$s = \frac{1}{2}(u + v)t$	$\theta = \frac{1}{2}(\omega_0 + \omega)t$
$s = ut + \frac{1}{2}at^2$	$\theta = \omega_0 t + \frac{1}{2}\alpha t^2$
$v^2 = u^2 + 2as$	$\omega^2 = \omega_0{}^2 + 2\alpha\theta$
$s = vt - \frac{1}{2}at^2$	$\theta = \omega t - \frac{1}{2}\alpha t^2$

ω = final angular velocity

ω_0 = initial angular velocity

α = angular acceleration

t = time

θ = angular distance

Worked example 1.12

Find the distance travelled when the final angular velocity is 14 rad/s, the angular acceleration is 3 rad/s^2 and the time of motion is 4 s.

θ (rad)	ω (rad/s)	t (s)	α (rad/s^2)
?	14	4	3

Use $\theta = \omega t - \frac{1}{2}\alpha t^2$

Substitute in given values to get:

$\theta = 14 \times 4 - \frac{1}{2} \times 3 \times 4^2 = 56 - 24 = 32$ rad

Worked example 1.13

Find the distance travelled when the initial angular velocity is 2 rad/s, the final angular velocity is 14 rad/s, and the time of motion is 4 s.

θ (rad)	ω_0 (rad/s)	ω (rad/s)	t (s)
?	2	14	4

Use $\theta = ½(\omega_0 + \omega)t$

Substitute in given values to get:

$\theta = ½(2 + 14) \times 4 = ½ \times 16 \times 4 = 32$ rad

Summary

- Angular displacement is the distance an object travels on a circular path and is often measured in radians. Tangential displacement is the distance an object moves in a straight line. Thus, an object that moves round a complete circle will have an angular displacement of 2π radians but a tangential displacement of 0 m (see Figure 1.36).
- Angular velocity is the angular displacement in a given unit of time. It is often measured in radians per second. Tangential velocity is the linear distance moved in a given unit of time and is often measured in metres per second.
- You can express angles in terms of revolutions, radians and degrees. For example, 1 revolution is the same as 2π radians or 360° (see Figure 1.37).
- Angular acceleration is the change in angular velocity per unit time. It is angular acceleration that contributes to the centripetal force which enables objects to move in circular paths. It is often measured in rad/s^2.
- Tangential displacement and angular displacement are related by the equation:

  tangential displacement = radius of path × angular displacement
- Tangential velocity and angular velocity are related by the equation

  tangential velocity = radius of path × angular velocity
- Angular acceleration and angular velocity are related by the equation:

$$\text{angular acceleration} = \frac{\text{angular velocity}}{\text{time}}$$

- The equations of motion with constant angular acceleration are related to the equations of motion with constant linear acceleration as shown in the table on page 28.

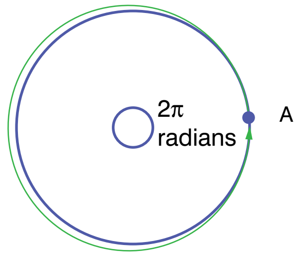


Figure 1.36

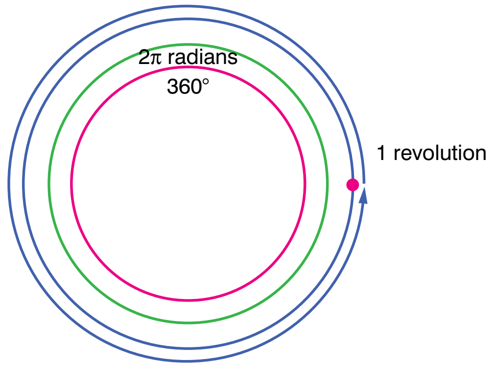


Figure 1.37

Review questions

1. a) Define angular displacement and gives its units.
   b) Define tangential displacement and give its units.
   c) Explain why an object that moves round a complete circle will have an angular displacement of 2π radians but a tangential displacement of 0 m.
2. a) Define angular velocity and give its units.
   b) Define tangential velocity and give its units.
   c) Find the tangential velocity of an object moving in a path of radius 2 m with an angular velocity of 3 rad/s.

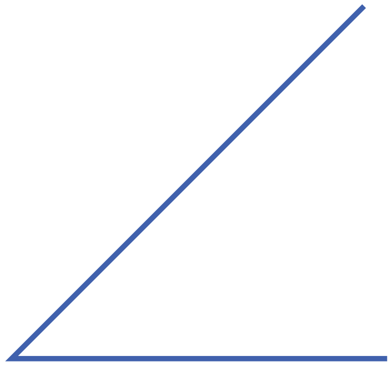


Figure 1.38

3. a) Give the three ways of expressing angles.
   b) Express the angle shown in Figure 1.38 in three different ways.
4. a) Define angular acceleration and give its units.
   b) How does angular acceleration enable objects to move in circular paths?
5. a) State the equations of motion with constant angular acceleration.
   b) Find the final angular velocity when the initial angular velocity is 5 rad/s, the angular acceleration is 2 rad/s and the time of motion is 10 s.
   c) Find the distance travelled when the final angular velocity is 20 rad/s, the angular acceleration is 2 rad/s and the time of motion is 5 s.
6. A satellite is in orbit 35600 km above the surface of the Earth. Its angular velocity is 7.27×10^{-5} rad/s. What is the velocity of the satellite? (The radius of the Earth is 6400 km.)
7. Astronauts in training are subjected to extreme acceleration forces by the centripetal forces in a giant centrifuge. The radius of the centrifuge is approximately 5 m.
   a) Calculate the approximate centripetal force on an astronaut of mass 80 kg if the centrifuge rotates once every 2 s.
   b) Approximately how many times greater than the astronauts weight is this force?

1.3 Rotational dynamics

By the end of this section you should be able to:

- Define the moment of inertia of a point mass.
- Define rotational kinetic energy of a body.
- Solve simple problems relating to moment of inertia and rotational kinetic energy.
- Define the term torque.
- Identify the SI unit of torque, which is N m, which is not the same as a joule.
- Express torque in terms of moment of inertia and angular acceleration.
- Derive an expression for the work done by the torque.

- Use the formula $W = \tau\theta$ to solve problems related to work done by torque.
- Define the angular momentum of a particle of mass *m* and write its SI unit.
- State the law of conservation of angular momentum.
- Solve problems using the law of conservation of angular momentum.
- State the first and second conditions of equilibrium.
- Solve problems related to conditions of equilibrium.
- Define the term centre of mass (centre of gravity) of a solid body.
- Determine the centre of gravity using a plumb-line method.
- Define the terms: stable, unstable and neutral equilibrium.

The moment of inertia of a point mass

The **moment of inertia** of a body is a measure of the manner in which the mass of that body is distributed in relation to the axis about which that body is rotating. It is dependent on the:

- mass of the body
- size of the body
- shape of the body
- which axis is being considered.

A spinning flywheel possesses **kinetic energy**, but how much? The expression $KE = \frac{1}{2}mv^2$ applies here, but the difficulty is that different parts of the flywheel are moving at different speeds – the regions further from the axis of rotation are going faster.

With linear motion, the kinetic energy is determined solely by the mass of the body and its speed. With the flywheel, the mass and the angular velocity are important, but there is now a third factor – how that mass is distributed in relation to the axis.

Consider two wheels each of mass *M*, but one is made in the form of a uniform disc whereas the other consists of a ring of the same radius *R* fixed to the axle by very light spokes rather like a bicycle wheel (Figure 1.39).

Both are spinning with the same angular velocity, ω. We can easily calculate the kinetic energy stored in the second flywheel because all its mass is at the rim.

You know that for a point mass going steadily in a circle of radius *r* at a linear speed *v*, that speed is related to its angular speed by:

$$\omega = \frac{v}{r} \text{ so } v = r\omega$$

KEY WORDS

moment of inertia *a measure of the distribution of the mass of a body in relation to its axis of rotation*

kinetic energy *the work needed to accelerate a body of a given mass from rest to its current velocity*

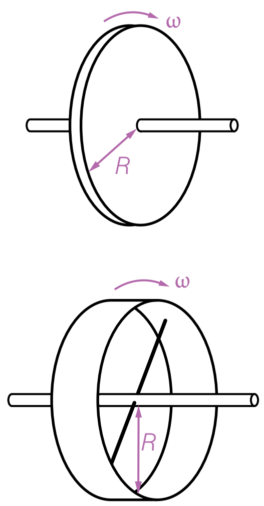


Figure 1.39 *Two wheels of the same mass but different construction.*

This flywheel has an angular speed ω, and since the whole of its mass is going in a circle of radius R it is all travelling at a linear speed $v = R\omega$. The kinetic energy, $\frac{1}{2}mv^2$, here will total $\frac{1}{2}MR^2\omega^2$.

What can we say about the first flywheel? Its total kinetic energy will be less, because most of its mass is moving at a speed slower than v. How can we go further than that?

A way forward is to think of the wheel as consisting of a large number of separate particles. There is no need to relate them to the individual atoms of the metal – we are just imagining it to be made up of a huge number of very small bits.

One of these bits, of mass m and distance r out from the axis, will have its share of the kinetic energy given by $\frac{1}{2}mv^2$ which can be expressed as $\frac{1}{2}mr^2\omega^2$ (since $v = r\omega$).

Activity 1.18: Rotational inertia

Place two different sized rolls of adding machine tape or other rolled paper on a dowel. Attach heavy clips to the rolls and hold them so that they cannot unwind. Release the rolls at the same time and note which unrolls most rapidly. Which roll has the greatest rotational inertia?

Rotational kinetic energy of a body

The total kinetic energy of the whole flywheel is just the sum of every particle in it. Those particles have different speeds v, but every one has the same angular velocity ω.

Adding all those kinetic energies, and denoting each particle with a subscript 1, 2, 3... and on for ever, we get:

Total kinetic energy $= \frac{1}{2}m_1r_1^{\,2}\omega^2 + \frac{1}{2}m_2r_2^{\,2}\omega^2 + \frac{1}{2}m_3r_3^{\,2}\omega^2 + \ldots$

We can rewrite this as:

$\frac{1}{2}(m_1r_1^{\,2} + m_2r_2^{\,2} + m_3r_3^{\,2})\omega^2 + \ldots$

The expression in the bracket is the rotational equivalent of mass in the expression of linear kinetic energy ($\frac{1}{2}mv^2$), and we call it the body's moment of inertia, I. Its units will be kg m^2.

We can simplify it if we replace all those separate values of r^2 by the *average* value of r^2 for *all* the particles in that body. It then becomes the total mass M of the body (which can be measured with an ordinary balance) multiplied by the average value of r^2 for all the particles (which can be calculated by a geometrical exercise for various shapes of body).

The moment of inertia of a body is its rotational inertia. It is not a constant for that body, because it depends on the axis chosen for it to rotate around. It may be defined for a given body about a given axis as the sum of mr^2 for every particle in that body (where m is the mass of the particle and r is its distance from the axis.

For a rotating body, its kinetic energy (KE) = $\frac{1}{2}I\omega^2$

Worked example 1.14

Find the moment of inertia of a point mass of 0.001 kg at a perpendicular distance of 2 m from its axis of rotation.

I (kg m²)	M (kg)	R (m)
?	0.001	2

Note that the perpendicular distance from a given axis of rotation is specified. This is the value you need for R in the equation.

Use $I = MR^2$

Substitute given values: $I = 0.001 \times 2^2 = 0.004$ kg m²

Worked example 1.15

Find the moment of inertia of a disc of radius 0.25 m and mass 0.5 kg.

I (kg m²)	M (kg)	R (m)
?	0.5	0.25

Use $I = \frac{1}{2}MR^2$

Substitute in given values: $I = \frac{1}{2} \times 0.5 \times 0.25^2 = 0.015625$ kg m²

Worked example 1.16

Find the moment of inertia of a sphere of mass 0.5 kg and radius 0.15 m.

I (kg m²)	M (kg)	R (m)
?	0.5	0.55

Use $I = \frac{2}{5}MR^2$

Substitute in given values: $I = \frac{2}{5} \times 0.5 \times 0.15^2 = 0.0045$ kg m²

Worked example 1.17

Find the kinetic energy of a rotating body with moment of inertia 0.004 kg m² and angular velocity of 0.5 rad/s.

KE (J)	I (kg m²)	ω (rad/s)
?	0.004	0.55

Use KE = $\frac{1}{2}I\omega^2$

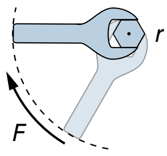
Substitute in given values: KE = $\frac{1}{2} \times 0.004 \times 0.5^2 = 0.0005$ J

Torque

A **torque** is a turning effect. It is the total moment acting on that body about the axis of rotation, and is measured by multiplying the force by its perpendicular distance from the axis.

torque = force × perpendicular distance

$$\tau = F \times r_{\text{perpendicular}}$$



The linear equivalent of a torque is a force.

The SI unit of torque is N m. This is not the same as a joule!

Activity 1.19: Comparing linear and rotational kinetic energy

Make sure that you understand why the expression for kinetic energy of a rotating body is $\frac{1}{2}I\omega^2$ by comparing it with the expression for kinetic energy of a body moving with linear motion.

Some moments of inertia

A point mass, M: $I = MR^2$

A disc of mass M and radius R (as with the first flywheel): $I = \frac{1}{2}MR^2$

A sphere of mass M and radius R rotating on its axis: $I = \frac{2}{5}MR^2$

KEY WORDS

torque *the tendency of a force to rotate an object about its axis of rotation*

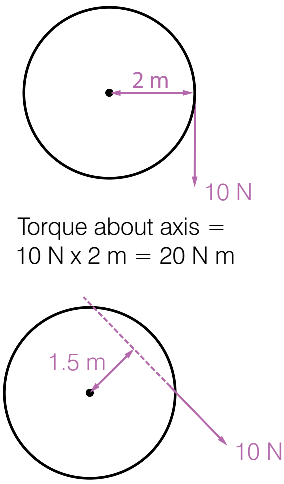


***Figure 1.40** The value of the torque depends on the distance from the axis.*

Activity 1.20: Distance in the torque definition

In a small group, discuss why distance is part of the definition of a torque.

Worked example 1.18

Find the work done when a torque of 5 N m moves through an angle of π radians.

W (J)	τ **(N m)**	θ **(rad)**
?	5	π

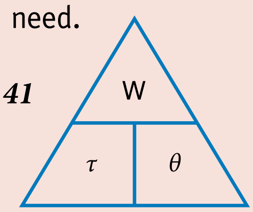
Use $W = \tau\,\theta$

Substitute given values

$W = 5\pi = 5\pi$ J

Remember that you can use this triangle to rearrange the formula. Cover up the term you are trying to find and the triangle will give you the form you need.

Figure 1.41



Worked example 1.19

Find the torque when the work done to move through an angle of π radians is 10 J.

W (J)	τ **(N m)**	θ **(rad)**
10	?	π

Use $W = \tau\,\theta$

Rearrange it so that torque is on the left-hand side.

$$\tau = \frac{W}{\theta}$$

Substitute given values

$$\tau = \frac{10}{\pi} \text{ N m}$$

Torque in terms of moment of inertia and angular acceleration

Since the linear equivalent of a torque, a force, can be found using Newton's second law:

force = mass × acceleration

you can see that, using the same law,

torque = moment of inertia × angular acceleration

If we give torque the symbol τ, then in symbols this equation is written as:

$\tau = I\alpha$

Work done by the torque

In linear terms the work done by a force is Fs, the force (F) multiplied by the distance (s) it moves in that direction. Likewise when a torque turns a body, the work it does = $\tau\,\theta$, the torque multiplied by the angle it turns through.

Angular momentum of a particle and its SI unit

The linear momentum of a body of mass m moving with a velocity v is defined to be mv. Likewise the **angular momentum** of a body of moment of inertia I rotating at an angular velocity ω is defined to be $I\omega$. Both linear momentum and angular momentum are vector quantities – they have magnitude and direction. Its units are N m s.

The law of conservation of angular momentum

The effect of an unbalanced force on a body is to cause its momentum to change. By Newton's second law of motion the momentum changes at a rate that is proportional to the magnitude of that force, and this leads to $F = ma$.

Similarly, the effect of an unbalanced torque on a body that can rotate is to cause its angular momentum to change, at a rate which is proportional to the magnitude of the torque. This leads to:

$\tau = I\alpha$ (the moment of inertia multiplied by the angular acceleration).

Just as linear momentum is conserved in the absence of a force, so is the angular momentum in the absence of a torque. The conservation of angular momentum says:

- if no resultant torque is acting, the angular momentum of a body cannot change.

This principle is demonstrated readily by a spinning skater, as shown in Figure 1.42. An ice skater holds her arms outstretched to either side and starts to spin round as fast as she can. She then folds her arms and her rate of rotation rises even further. By bringing more of her mass in closer to the axis of rotation, she has reduced her moment of inertia. Since no external torque is acting, $I\omega$ has to stay constant. Here I is made less, so ω has to increase.

Activity 1.21: Experience the conservation of angular momentum

If you have access to a chair that can swivel round freely, you can do this for yourself.

Hold a large mass in each hand and extend your arms fully to either side. Sit in the chair and get your friends to turn the chair round as fast as they can. Tell them to stop turning, and bring the masses close in to your body. Your moment of inertia has fallen so you should spin faster. Your angular momentum is conserved.

Your rotational kinetic energy is $\frac{1}{2} I\omega^2$. This can be written as $\frac{1}{2}(I\omega)\omega$. As you bring your arms in, the $(I\omega)$ term stays constant. You are spinning faster, though, so that extra factor of ω means you have gained in rotational kinetic energy.

Where has it come from? If you have tried this, you may well have felt the answer. As those two masses were pulled into a tighter circle, you had to provide the increasing centripetal force to achieve this. You had to exert a force, and had to move that force through a distance. In other words, you did some work.

Therefore you had to release some of the chemical energy in your food – and that is where the extra kinetic energy came from.

Worked example 1.20

No external force acts on a skater. Her moment of inertia is initially 60 kg m². Her angular velocity is 0.2 rad s⁻¹ at the beginning of a spin. She brings in her arms and her angular velocity increases so that her moment of inertia decreases to 50 kg m². Find her final angular velocity.

initial angular momentum = final angular momentum

I (kg m²)	ω (rad/s)
60	0.2

I (kg m²)	ω (rad/s)
50	?

angular momentum at end of spin = angular momentum at beginning of spin

$60 \times 0.2 = 50 \times$ final angular velocity

$$\text{final angular velocity} = \frac{60 \times 0.2}{50}$$

0.24 rad/s

Figure 1.42 *Spinning skater.*

Activity 1.22: Applications of the principle of conservation of angular momentum

In a small group, make a list of applications of the principle of conservation of angular momentum. To start you off, think about other sports where spinning is required.

Activity 1.23: Comparing linear and rotational motion

In a small group, make a poster to compare linear and rotational motion. You should include all the information you have learnt in this unit, and rotational equivalents of Newton's first and second laws.

KEY WORDS

equilibrium *a stable situation in which forces cancel one another out*

This table summarises the relationships between linear and angular quantities.

Linear	Angular
$P = mv$	$L = I\omega$
$F = ma$	$\tau = I\alpha$
$F = \frac{\Delta p}{\Delta t}$	$\tau = \frac{\Delta L}{\Delta t}$
$\Delta p = 0$	$\Delta L = 0$
$a = \frac{\Delta v}{\Delta t}$	$\alpha = \frac{\Delta \omega}{\Delta t}$

The conditions of equilibrium

There are two conditions that must be satisfied if a body is to be in equilibrium.

1 The forces acting on it must sum to zero.

2 The turning effects of the forces must sum to zero.

Worked example 1.21

Explain why these systems are in equilibrium.

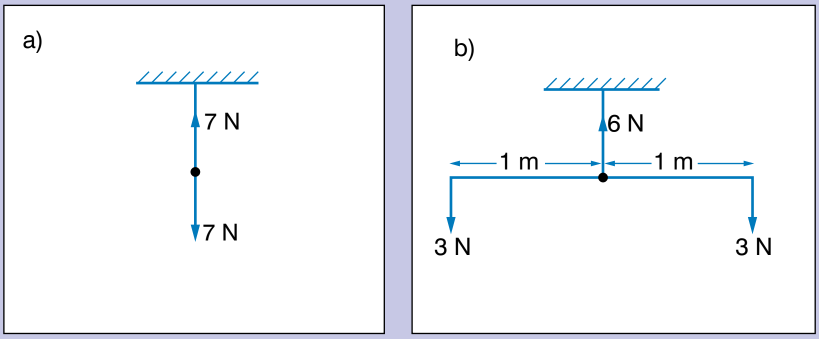


Figure 1.43

In a) there is no turning effect so condition 2 above is satisfied. The forces are 7 N upwards and 7 N downwards so the forces acting sum to zero, satisfying condition 1 above. So a) is in equilibrium.

In b) there would be a turning effect but one is 3 N m anticlockwise and the other is 3 N m clockwise to they sum to zero so condition 2 above is satisfied. The forces are 6 N upwards and a total of 6 N downwards so these sum to zero so condition 1 above is satisfied. So b) is in equilibrium.

The centre of mass (centre of gravity) of a solid body

The **centre of mass** of a solid body is the point at which the body's whole mass can be considered to be concentrated for the purpose of calculations. In a solid body, the position of its centre of mass is fixed in relation to the object (but not necessarily in contact with it). The centre of mass is often called the centre of gravity but this is only true in a system where the gravitational forces are uniform, such as on Earth.

Activity 1.24: Determine the centre of gravity using a plumb-line method

Repeat Activity 1.13 (see page 19) but this time, rather than using a piece of thick card, use an object made from modelling clay.

DID YOU KNOW?

Engineers try to design a sports car's centre of mass as low as possible to make the car handle better. When high jumpers perform a 'Fosbury Flop', they bend their body in such a way that it is possible for the jumper to clear the bar while his or her centre of mass does not.

Figure 1.44 *These spheres will return to their original position when they are displaced.*

Stable, unstable and neutral equilibrium

An object is in **stable equilibrium** if, when it is slightly displaced, it returns to its original position.

An object is in **unstable equilibrium** if, when it is slightly displaced, it moves further away from its original position.

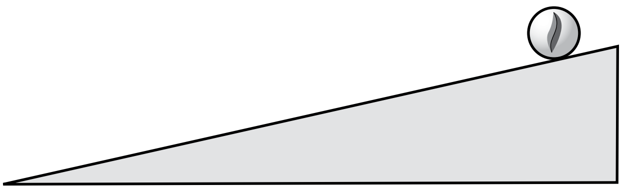


Figure 1.45 *If this marble is displaced it will run down the ramp and not return to its original position.*

An object is in **neutral equilibrium** if, when it is slightly displaced, the system does not necessarily return to its original position but neither does it move further away. For example, if you kick a football along the ground it will roll a little way and then stop at another spot. The kick changes its position but not its stability.

KEY WORDS

stable equilibrium *the tendency of an object, if it is displaced, to return to its original position*

unstable equilibrium *the tendency of an object, if it is displaced, to keep moving and not to return to its original position*

neutral equilibrium *the tendency of an object, if it is displaced, to neither return to its original position nor to move farther from it*

Figure 1.46 *This is neutral equilibrium.*

DID YOU KNOW?

The concept of centre of mass was first introduced by the ancient Greek mathematician, physicist, and engineer Archimedes. He showed that the torque exerted on a lever by weights resting at various points along the lever is the same as what it would be if all of the weights were moved to a single point — their centre of mass. In work on floating bodies he demonstrated that the orientation of a floating object is the one that makes its centre of mass as low as possible.

DID YOU KNOW?

The centre of mass on an aircraft significantly affects the stability of the aircraft. To ensure the aircraft is safe to fly, the centre of mass must fall within specified limits.

Summary

- The moment of inertia of a point mass, M, is $I = MR^2$.
- The moment of inertia for a disc of mass M and radius R is: $I = \frac{1}{2} MR^2$.
- The moment of inertia for a sphere of mass M and radius R rotating on its axis is: $I = \frac{2}{5} MR^2$.
- The rotational kinetic energy of a body is $\frac{1}{2}I\omega^2$.
- You can use the above formulae for simple problems relating to moment of inertia and rotational kinetic energy.
- A torque is a turning effect. It is the total moment acting on that body about the axis of rotation, and is measured by multiplying the force by its perpendicular distance from the axis. $\tau = F \times r_{perp}$. Its unit is N m, which is not the same as a joule.
- torque = moment of inertia × angular acceleration
- If we give torque the symbol τ, then in symbols this equation is written as:

  $\tau = I\alpha$
- When a torque turns a body the work it does = $\tau\,\theta$, the torque multiplied by the angle it turns. You can use this formula $W = \tau\theta$ to solve problems related to work done by torque.
- The angular momentum of a body of moment of inertia I rotating at an angular velocity ω is defined to be $I\omega$. Its units are N m s.
- The conservation of angular momentum says: if no resultant torque is acting, the angular momentum of a body cannot change.
- You can use this law to solve problems.
- There are two conditions that must be satisfied if a body is to be in equilibrium.

  1 The forces acting on it must sum to zero.

  2 The turning effects of the forces must sum to zero.
- You can solve problems using the conditions of equilibrium.
- The centre of mass of a solid body is the point at which the body's whole mass can be considered to be concentrated for the purpose of calculations. You can determine the centre of gravity using a plumb-line method.
- An object is in stable equilibrium if, when it is slightly displaced, it returns to its original position.
- An object is in unstable equilibrium if, when it is slightly displaced, it moves further away from its original position.

- An object is in neutral equilibrium if, when it is slightly displaced, the system does not necessarily return to its original position but neither does it move further away.

Review questions

1. Define the moment of inertia of a point mass.
2. Define rotational kinetic energy of a body.
3. a) Find the moment of inertia of a point mass of 0.005 g at a perpendicular distance of 3 m from its axis of rotation.
   b) Find the moment of inertia of a sphere of mass 0.3 kg and radius 0.6 m.
   c) Find the kinetic energy of a rotating body with moment of inertia 0.003 kg m^2 and angular velocity of 0.6 rad s^{-1}.
4. Define the term torque and identify its SI unit.
5. Express torque in terms of moment of inertia and angular acceleration.
6. Derive an expression for the work done by the torque.
7. a) Find the work done when a torque of 3.5 N m moves through an angle of $\frac{1}{2}\pi$ radians.
   b) Find the torque when the work done to move through an angle of $\frac{1}{4}\pi$ radians is 3 J.
8. Define the angular momentum of a particle of mass *m* and write its SI unit.
9. State the law of conservation of angular momentum.
10. Give examples of uses of the law of conservation of angular momentum.
11. State the first and second conditions of equilibrium.
12. Explain why the system in Figure 1.47 is in equilibrium?

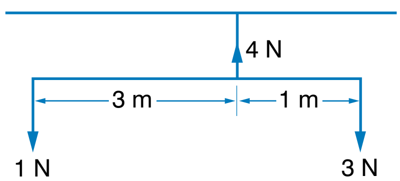


Figure 1.47

13. Define the term centre of mass (centre of gravity) of a solid body.
14. Explain how you can determine the centre of gravity using a plumb-line method.
15. Define the terms stable, unstable and neutral equilibrium.

1.4 Newton's law of universal gravitation

By the end of this section you should be able to:

- State Newton's law of universal gravitation.
- Determine the magnitude of the force of attraction between two masses separated by a distance r.
- Calculate the value of g at any distance above the surface of the Earth.
- State Kepler's laws of planetary motion.
- Use Kepler's laws of planetary motion to determine the period of any planet.
- Differentiate between the orbital and escape velocity of a satellite.
- Determine the period of a satellite around a planet.
- Calculate the orbital and escape velocity of a satellite.
- Describe the period, position and function of a geostationary satellite.

Newton's law of universal gravitation

To simplify matters we shall consider that the planets go round the Sun in circular orbits, which is nearly the case. There is no force opposing their motion, which is why they just keep going.

There has to be a force pulling them towards the Sun, though. Otherwise they would keep moving in a straight line. If the planet has a mass m, is travelling at a speed v and follows an orbit of radius r, the magnitude of the force straight towards the Sun has to be $\frac{mv^2}{r}$.

There is nothing but space between the planet and the Sun. What is producing a force that size? The answer is gravitation. This is a force of attraction that exists between any two lumps of matter, but which becomes noticeable only if at least one of them is of astronomical size.

Based on observations of the path of the Moon, Newton proposed a formula to describe how this force must behave. In doing so, he imagined that the laws of nature that applied to objects on Earth also applied to heavenly bodies. At the time this was a very daring idea.

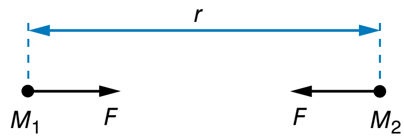


Figure 1.48 *A gravitational force attracting two masses.*

The law has since been confirmed by incredibly sensitive experiments carried out between a pair of masses on Earth, and is known as Newton's universal law of gravitation.

If two masses M_1 and M_2 are a distance r apart, Newton claimed that the gravitational force F between them was proportional to each of the masses, and decreased as they moved apart by an inverse square relationship – move three times as far apart and the force drops to one ninth, for example.

Putting this together, we get:

$$F = \frac{GM_1M_2}{r^2}$$

G is a constant for all matter everywhere, and is called the gravitational constant. You should be able to see that it will have units of N m^2 kg^{-2}.

Its value has been measured to be 6.67×10^{-11} N m^2 kg^{-2}. There is no need to try to remember that, but its small size does illustrate the weakness of the force.

Worked example 1.22

Find the gravitational force between the Earth and the Moon. The mass of the Moon is 7.35×10^{22} kg, the mass of the Earth is 5.98×10^{24} kg and the distance between them is approximately 3×10^6 km.

F (N)	**G (N m^2 kg^{-2})**	**M_1 (kg)**	**M_2 (kg)**	**r (m)**
?	6.67×10^{11}	7.35×10^{22}	5.98×10^{24}	3×10^9

Use $F = \dfrac{GM_1M_2}{r^2}$

Substitute in given values and value for G.

$$F = \frac{6.67 \times 10^{-11}\ \text{N m}^2\ \text{kg}^{-2} \times 7.35 \times 10^{22}\ \text{kg} \times 5.98 \times 10^{24}\ \text{kg}}{(3 \times 10^9\ \text{km})^2}$$

$$F = \frac{2.93 \times 10^{37}}{9 \times 10^{18}} = 3.257 \times 10^{18}\ \text{N}$$

Finding the value of g at any distance above the surface of the Earth

The value of g varies a little above the surface of the Earth. The law of universal gravitation tells us that the force on a body acted upon by Earth's gravity is given by:

$$F = G\frac{m_1m_2}{r^2} = \left(G\frac{m_1}{r^2}\right)m_2$$

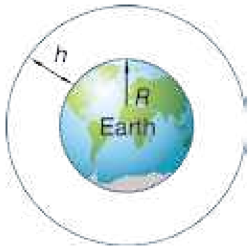


Figure 1.49 To calculate g *at a height* h *above the Earth, use* R + h *instead of* r *in the formula.*

where r is the distance between the centre of the Earth and the body, and m_1 is the mass of the Earth and m_2 is the mass of the body.

Newton's second law, $F = ma$, where m is mass and a is acceleration, tells us that:

$F = m_2 g$

Comparing the two formulae you can see that:

$$g = G\frac{m_1}{r^2}$$

To calculate the value of g at a distance h above the surface of the Earth, you need to substitute the value of the radius of the Earth R, + h, for r in the above formula (see Figure 1.49).

Worked example 1.23

Find the value of g at a distance of 2 km above the surface of the Earth.

g (m/s^2)	G ($N\ m^2\ kg^{-2}$)	m_1 (kg)	r_2 (m)
?	6.67×10^{-11}	5.98×10^{24}	$6.4 \times 10^6 + 2000$

Use $g = \frac{G\,m_1}{r^2}$

$$g = \frac{6.67 \times 10^{-11} \times 5.98 \times 10^{24}}{(6.4 \times 10^6 + 2 \times 10^3)^2}$$

$$g = \frac{3.98 \times 10^{14}}{(6\ 402\ 000)^2}$$

$$g = \frac{3.98 \times 10^{14}}{4.099 \times 10^{13}}$$

$g = 9.71\ m/s^2$

Kepler's laws of planetary motion

Kepler's laws are experimental laws that describe the rotation of satellites about their parent body. They apply to all the satellites around the Earth, but Kepler's data were collected in the late 1500s – before the telescope had been invented – by closely observing the paths of the planets as they went around the Sun.

Kepler began with many sheets of figures, and amazingly succeeded in spotting three unexpected patterns to them all. These are now known as **Kepler's laws**:

1 Each planet moves in a path called an ellipse, with the Sun at one focus.

KEY WORDS

Kepler's laws *describe the motion of a planet around the Sun*

2 The line that joins the Sun to the orbiting planet sweeps out equal areas in equal times.

3 The square of the time it takes the planet to go round the Sun (that is, the square of its year) is proportional to the cube of its average distance from the Sun.

This law can be written in equation form as

$$T^2 = \left[\frac{4\pi}{GM}\right] a^3$$

where T is the time it takes the planet to go round the Sun, G is the gravitational constant, M is the mass of the planet and a is the mean distance between the planet and the Sun.

We will look at the first two laws. Figure 1.50 illustrates what they are saying. The two shaded areas are equal.

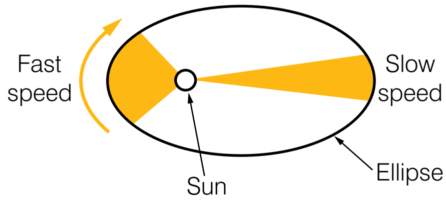


Figure 1.50 *Kepler's laws 1 and 2.*

The second law follows from the conservation of energy. As the planet gets closer to the Sun it speeds up, and as it climbs back to a more distant part of its orbit some of that kinetic energy is transferred into its extra gravitational potential energy.

The planets all have orbits round the Sun that are close to being circles. There are some natural satellites of the Sun that are very different, though – the comets. These are only visible to us as they pass quickly through the part of their orbit which is close to the Sun. Most of the time they are remote from the Sun, in the darkness of space and travelling more slowly.

The path of the Moon as it travels round the Earth is very nearly a circle, though there are times in its orbit when it is closer to us and so appears slightly larger in the sky.

Halley's Comet

We know that this has appeared to us in the sky every 75–76 years during recorded history, right back to the year 240 BC. It was last visible in 1986 and is next due in 2061. All the planets go round the Sun in the same direction. Halley's Comet goes the opposite way round, and its orbit is tilted at about 18° to that of the planets. This strongly suggests that it did not form when the planets formed, but passed by the Sun at a later time close enough to find itself captured by it.

The period, position and function of a geostationary satellite

The Earth nowadays has a large number of satellites in orbit round it. Most were made by humans, but the biggest one of them all is natural – the Moon.

Since the path of the Moon as it travels around the Earth is very nearly a circle, we can say that it has constant speed and is subject to a centripetal force

$$F = \frac{mv^2}{r}$$

where m is the mass of the Moon, v is its speed and r is its distance from the Earth.

If Earth has mass M, Newton's law of universal gravitation tells us that

$$F = \frac{GMm}{r^2}$$

So $\dfrac{mv^2}{r} = \dfrac{GMm}{r^2}$

$$v^2 = \frac{GMr}{r^2} = \frac{GM}{r}$$

$$v = \sqrt{\frac{GM}{r}}$$

This equation can be applied to all satellites. This means that the smaller the distance from Earth, the greater the speed of the satellite.

Worked example 1.24

Find the speed of the Moon if the mass of the Earth is 6×10^{24} kg and the distance from the Moon to Earth is 4×10^{8} m.

v (m/s)	G (N m^2 kg^{-2})	M (kg)	r (m)
?	6.67×10^{-11}	6×10^{24}	4×10^{8}

Use $v = \sqrt{\frac{GM}{r}}$

$$v = \sqrt{\frac{6.67 \times 10^{11} \times 6 \times 10^{24}}{4 \times 10^{8}}}$$

$$v = \sqrt{\frac{4.002 \times 10^{36}}{4 \text{ x } 10^{8}}}$$

$$v = \sqrt{1.0005 \text{ x } 10^{28}}$$

$$v = 1.00025 \text{ x } 10^{14}$$

Geostationary satellites orbits are above the equator and they are going round with the same angular velocity as the Earth. This means that they have the same period as the Earth. They are always in the same spot in the sky and so they are ideal for communications purposes and for satellite navigation systems.

KEY WORDS

geostationary satellite *a satellite that orbits the equator with the same angular velocity as the Earth*

Orbital velocity of a satellite

Other satellites have orbits that are always moving relative to the Earth. Many of these are used to look down on the Earth beneath. Some are military spy satellites, but others observe the Earth – the weather over the whole globe, the temperatures of the oceans and so on. One is even used to provide a platform for a telescope to study the rest of the universe clearly without having to peer through the Earth's atmosphere.

The satellites are all launched from sites as close to the equator as possible, and certainly not from polar regions. The Earth is spinning on its axis: the nearer you are to the equator, the greater the distance you must travel to go round once in a day, and therefore the faster your speed is due to this motion. If you wanted to throw a ball at as great a speed as you could, propelling it forward from a fast-moving car would give you a good start. Satellites launched by humans similarly take advantage of the Earth's motion.

Activity 1.25: Verify Kepler's third law

On page 27, the orbital period of five planets is given. Use this data to verify Kepler's third law.

Activity 1.26: Use of geostationary satellites

Research some uses of geostationary satellites.

A satellite is not always sent into its final orbit in one step. Sometimes it will be launched so it goes into a temporary 'parking orbit'. From there a second carefully controlled rocket can be fired to lift it into its permanent orbit.

All satellites are in orbits that are distant from Earth. There is no mathematical reason why one should not orbit just above the Earth's surface, but the inconvenience of it is not the main reason why nobody does that. The satellites are deliberately positioned so that they are clear of the Earth's atmosphere – this means there is no resistance force to oppose the satellite's horizontal motion, otherwise an engine would be needed to cancel out such a force. Each satellite moves at a particular velocity in its orbit – this is its orbital velocity and is calculated by using the equation $\omega = \frac{2\pi}{T}$ where ω is its angular velocity and T is its period.

Activity 1.27: Calculate the orbital velocity of a satellite

Use $\omega = \frac{2\pi}{T}$ to calculate the orbital velocity of a geostationary satellite.

Escape velocity of a satellite

If a satellite is launched vertically upward at a sufficiently large velocity, it will be able to climb right out of the Earth's potential well and escape completely.

We can calculate this speed. Suppose the satellite has a mass m. At the Earth's surface the potential is $\frac{GM}{R}$ in J kg^{-1},

where M and R are the Earth's mass and radius. Thefore the satellite on the Earth's surface will have a potential energy of $\frac{GMm}{R}$ in joules.

When completely clear of the Earth, its potential energy will be zero. Therefore we must raise its potential energy by $\frac{GMm}{R}$ in joules. If the satellite is launched with a speed of v upwards, it will have an amount of kinetic energy given by $\frac{1}{2}mv^2$. If this kinetic energy is enough to supply what is needed, it can escape.

Thus the minimum escape velocity is given by:

$$\frac{1}{2}mv^2 = \frac{GMm}{R}$$

whence $v = \sqrt{\frac{2GM}{R}}$

The expression for the escape speed can be made even simpler. If we substitute $M = \frac{gR^2}{G}$ in the above, we get:

escape velocity $v = \sqrt{(2gR)}$

Worked example 1.25

Find the escape velocity for a satellite leaving Earth. The radius of the Earth is 6.4×10^6 m and g is 9.81 N kg^{-1}

v (m/s)	g (N kg^{-1})	R (m)
?	9.81	6.4×10^6

Use $v = \sqrt{(2gR)}$

$v = \sqrt{(2 \times 9.81 \times 6.4 \times 10^6)}$

$v = \sqrt{(125\ 568\ 000)}$

$v = 11\ 205.71$ m/s

$v = 11$ km/s

Summary

- Newton's law of universal gravitation states that if two masses M_1 and M_2 are a distance r apart, the gravitational force F between them is proportional to each of the masses, and decreases as they move apart by an inverse square relationship:

  $$F = \frac{GM_1M_2}{r^2}$$

- G is a constant for all matter everywhere, and is called the gravitational constant. You should be able to see that it will have units of N m^2 kg^{-2}.
- You can use Newton's law of universal gravitation to determine the magnitude of the force of attraction between two masses separated by a distance r.
- You can calculate the value of g at any distance above the surface of the Earth using:

  $$g = G\frac{m_1}{r^2}$$

- To calculate the value of g at a distance h above the surface of the Earth, substitute the value of the radius of the Earth, $R + h$ for r in the above formula.
- Kepler's laws of planetary motion are:

  1. Each planet moves in a path called an ellipse, with the Sun at one focus.
  2. The line that joins the Sun to the orbiting planet sweeps out equal areas in equal times.
  3. The square of the time it takes the planet to go round the Sun (that is, the square of its year) is proportional to the cube of its average distance from the Sun.

- You can use Kepler's laws of planetary motion to determine the period of any planet.
- You can determine the period of a satellite around a planet using $T = \frac{2\pi}{\omega}$.
- Geostationary satellite orbits are above the equator and they are going round with the same angular velocity as the Earth. This means that they have the same period as the Earth. They are always in the same spot in the sky and so they are ideal for communications purposes and for satellite navigation systems.
- The orbital velocity of a satellite is the speed with which it goes round the Earth (or other planet). The escape velocity of a satellite is the speed it needs to escape the potential well of the Earth.

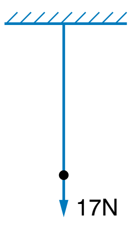
- You can calculate the orbital velocity of a satellite using $\omega = \frac{2\pi}{T}$ and the escape velocity of a satellite using $v = \sqrt{(2gR)}$.

Review questions

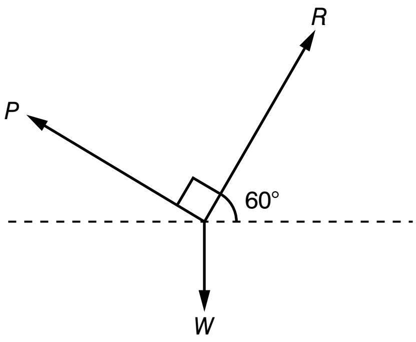
1. State Newton's law of universal gravitation.
2. Determine the magnitude of the force of attraction between Mercury and the Sun. They are approximately 5.8×10^{10} m apart. The mass of Mercury is about 3.3×10^{23} kg and the mass of the Sun is about 2×10^{30} kg. G is 6.67×10^{-11} N m^2 kg^{-2}.
3. Calculate the value of g at 1000 m above the surface of the Earth. The radius of the Earth is 6378.1 km.
4. State Kepler's laws of planetary motion.
5. Describe the period, position and function of a geostationary satellite.
6. Differentiate between orbital and escape velocity of a satellite.

End of unit questions

1. a) Define the term 'projectile'.
   b) Give some examples of projectiles.
2. a) Explain why the angle is important when launching projectiles.
   b) Find the range of a projectile launched at an angle of 45° with an initial velocity of 25 m/s.
   c) Explain why the vertical velocity of a projectile does not change.
3. a) State the centre of mass theorem.
   b) Give some practical applications of centre of mass.
4. Draw a table to compare the equations of motion with constant angular acceleration with the equations of motion with constant linear acceleration.
5. Find the distance travelled when the initial angular velocity is 3 rad/s, the final angular velocity is 25 rad/s , and the time of motion is 10 s.
6. Find the moment of inertia of a disc of radius 0.4 m and mass 0.75 kg.
7. Copy and complete this diagram so that the system is in equilibrium.



8. This system is in equilibrium. Write an expression linking P, R and W.



9. Determine the magnitude of the force of attraction between the Moon and the Earth. They are approximately 4×10^8 m apart. The mass of the Moon is 2×10^{24} kg and the mass of the Earth is 6×10^{24} kg.
10. Geostationary satellites are placed in orbits of radius 4.2×10^4 km. Use this information to deduce g at that height.
11. A climber of mass 80 kg is on a steep rock face. The force X that the rock exerts on the climber is at an angle of 50° to the vertical. Y, the other force on the climber, keeps him in equilibrium and is provided by a rope at an angle of 40° to the vertical.

    a) Draw a sketch to show the forces acting on the climber.

    b) From your sketch of the forces, sketch a triangle of forces to show equilibrium.

    c) Use your triangle of forces to find

       i) X, the force the rock face exerts on the climber

       ii) Y, the force provided by the rope.
12. Which of these answers is correct? Justify your answer.

    If the polar ice caps melt completely as a result of rising global temperatures, then

    a) the Earth will rotate faster

    b) the Earth will rotate slower

    c) there will be no change in the angular speed

    d) the duration of a day on the Earth will increase.

Electrostatics

Unit 2

Contents

Section	Learning competencies
2.1 Electric charge (page 50)	• State the law of conservation of charge. • Describe and explain the charging processes: charging by rubbing, conduction and induction. • Perform an experiment to charge an electroscope by conduction and by induction. • Describe the distribution of charge on a conductor of variable shape. • Explain how lightning is formed. • Describe the use of a lightning rod. • Describe how equipment works using electrostatics principles. • Describe hazards and uses of electrostatics.
2.2 Electric forces and fields (page 60)	• Define an electric field. • Represent diagrammatically the electric field lines around and between two points. • Distinguish between the electric field inside, outside and between surfaces of a spherical metallic conductor. • State Coulomb's law. • Compare Coulomb's law and Newton's law of universal gravitation. • Calculate the force acting on a charge due to two other charges placed on the same plane (line of action). • Calculate the force between three charges placed in a line. • Calculate the electric field strength at a point due to charges placed in a line and at right angles.
2.3 Electric potential (page 67)	• Define electric potential and its SI unit. • Distinguish between absolute potential and potential difference. • Show that 1 N/C = 1 V/m. • Explain equipotential lines and surfaces. • Draw equipotential lines and surfaces in an electric field. • Define the term electric potential energy.
2.4 Capacitors and capacitances (page 70)	• Describe the structure of a simple capacitor. • Define the term capacitance and its SI unit. • Apply the definition of capacitance to solve numerical problems. • Use the circuit symbol to represent a capacitor. • Explain the charging and discharging of a capacitor. • Define the term dielectric and explain what is meant by a dielectric material. • Identify combination of capacitors in series, parallel and series–parallel. • Explain the effect of inserting a dielectric in the gap between the plates of a parallel plate capacitor.

Contents

Section	Learning competencies
	• Derive an expression for the effective capacitance of capacitors connected in series and parallel. • Draw an electric circuit diagram for a simple capacitor, series and parallel connections of two or more capacitors using symbols. • Solve problems on combination of capacitors. • Define parallel plate capacitor. • Describe the factors that affect the capacitance of a parallel plate capacitor. • Calculate the capacitance of a parallel plate capacitor. • Find an expression for the electric potential energy stored in a capacitor. • Calculate the energy stored in a capacitor using one of three possible formulae. • State some uses of capacitors in everyday life.

KEY WORDS

electrostatics *the build up of electric charge on the surface of objects*

charge *an electric charge, positive, negative, or zero, found on the elementary particles of matter*

atoms *the smallest components of an element having the chemical properties of that element*

positive charge *having a deficiency of electrons*

negative charge *having a surplus of electrons*

2.1 Electric charge

By the end of this section you should be able to:

- State the law of conservation of charge.
- Describe and explain the charging processes: charging by rubbing, conduction and induction.
- Perform an experiment to charge an electroscope by conduction and by induction.
- Describe the distribution of charge on a conductor of variable shape.
- Explain how lightning is formed.
- Describe the use of a lightning rod.
- Describe how equipment works using electrostatics principles.
- Describe hazards and uses of electrostatics.

Electrostatics in everyday life

Have you ever felt a small electric shock when you have touched a metal door handle, or heard a crackling sound as you pull a woollen article of clothing over your head? Both these arise because of **electrostatics** – electric **charge** is transferred from one material to another. In this section you will learn more about this, its applications and its hazards.

Activity 2.1: Recalling other physical quantities that are conserved

In a small group, spend one minute writing a list of physical quantities that you have met that follow a conservation law. Hint: think back to Unit 1.

What is electric charge?

All objects are made up from tiny building blocks called atoms. Individual atoms are made up of particles that possess both mass and electric charge. A material such as Perspex is made up from a huge number of atoms, but although we are well aware of its mass, it does not seem to possess any charge at all. Why is this?

Every bit of mass in every atom adds up so that we feel its mass. In every single atom, however, the two sorts of electric charge cancel one another out: there is the same number of positive charges as negative ones, so we do not notice any overall charge. An 'uncharged' piece of Perspex contains a vast number of charges, but the numbers of the two kinds are equal. Charging it up involves upsetting the balance between positive and negative charges. Adding or removing a few more electrons would do this.

The elementary charge, $e = -1.6 \times 10^{-19}$ C, is the amount of charge on an electron (negative charge) or on a proton (positive charge). To find the total charge Q, you use $Q = N \times e$, where N is the number of electrons.

KEY WORDS

law of conservation of charge *the total electric charge of a system remains constant despite changes inside the system*

DID YOU KNOW?

The fact that some materials (such as amber) become charged after they are rubbed has been known for thousands of years. The Greek word for amber, *ἤλεκτρον* (electron), is the source of the word 'electricity'.

Conservation of charge

Electric charge cannot be created or destroyed, only transferred. The number of **positive** charges on an atom of a substance is the same as the number of **negative** charges on the atom. The atom is neutral. The charges from some atoms can be removed from the atom and transferred to another material in the process known as charging but the overall number of positive and negative charges does not change. This is the **law of conservation of charge**.

Like charges will repel each other (for example, two positive charges will repel each other, two negative charges will repel each other). Unlike charges will attract each other (so a positive charge will attract a negative charge and a negative charge will attract a positive charge).

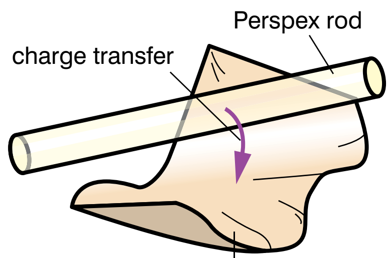


Figure 2.1 *Rubbing Perspex transfers charge from the surface of the Perspex to the material you are rubbing it with.*

Charging materials by rubbing

It is possible to charge some materials by rubbing them. When you rub a piece of Perspex, for example, some of the charge is transferred from the surface of the Perspex to the material you are using to rub it, and so the overall charge on the Perspex becomes unbalanced.

Activity 2.2: Testing how charged bodies attract or repel one another

Take a bar of Perspex and pivot it. One way is to suspend it by a nylon thread (cotton, if the slightest bit damp, may allow the charges to leak off too readily). Charge the suspended rod by rubbing it. Bring a second charged Perspex rod up to the first one, and you will see the first one swing away.

Now charge the bar of Perspex again. Bring a charged rod of Polythene up to the suspended rod, and you will see the suspended rod moving towards it. What does this tell you about the charges on the Perspex and the Polythene?

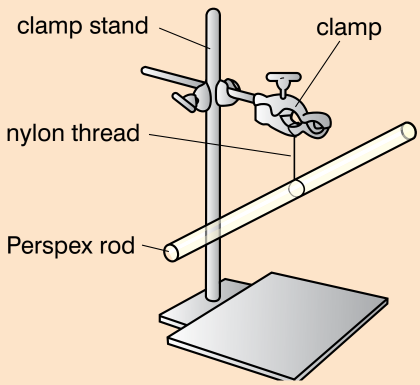


Figure 2.2 *Apparatus to test attraction or repulsion between charged bodies.*

Activity 2.3: The electrostatic attraction of water

Adjust a tap so a continuous but gentle stream of water is falling from it. Rub a plastic comb on your sleeve, and bring it up to the side of the water column. Describe and try to explain what happens.

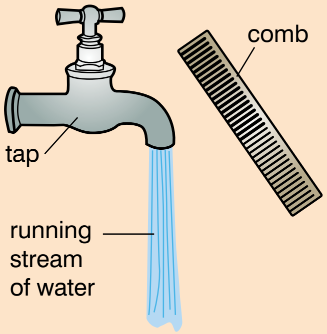


Figure 2.3 *The electrostatic attraction of water.*

Activity 2.5: Charging an electroscope by conduction

Charge a metal sphere. Bring it into contact with an electroscope. What happens? Draw a diagram to explain your results.

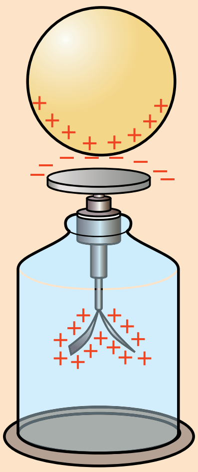


Figure 2.7 *Charging an electroscope by conduction.*

Activity 2.4: Hanging balloons using electrostatic attraction

Inflate a balloon and then briskly rub one side of it on your hair. Place the surface that you have rubbed towards a wall or door and release it when it appears to be sticking. What can you say about the nature of the surfaces to which the balloon sticks?

Figure 2.4 *Rubbing a balloon on hair.*

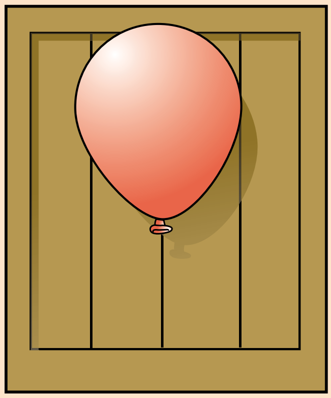


Figure 2.5 *A balloon sticking to a surface.*

Charging materials by conduction

If a charged object is brought into contact with an uncharged (neutral) object, then the neutral object will become charged by **conduction**. Charge flows from the charged object to the neutral object. (This is what happens when you feel an electric shock when you touch a metal door handle – the door handle has become charged and the charge travels by conduction to you, a neutral object!)

Figure 2.6 *Charging a person by conduction.*

Charging by induction

Figure 2.8 shows a different way of charging a body known as **induction**.

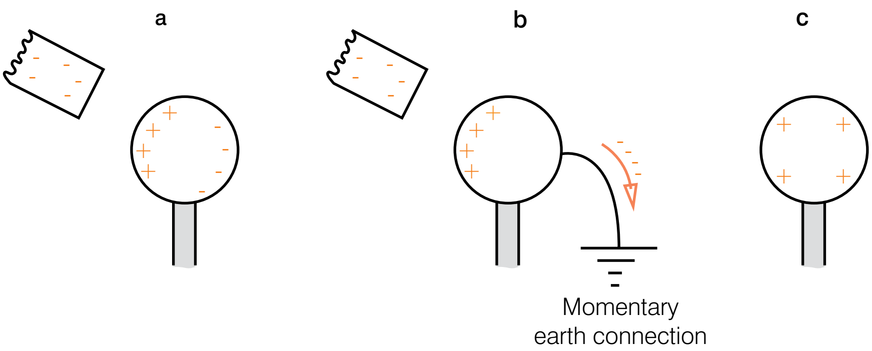


Figure 2.8 *Charging by induction.*

In Figure 2.8a, a negatively charged rod is brought near to a neutral metal sphere held on an insulating support. The negative charges in the metal sphere move to the side of the sphere furthest from the rod, leaving the side of the sphere nearest to the rod positively charged. In Figure 2.8b, while the rod is still nearby, you touch the sphere for a moment then let go. The negative charges now escape further, through you down to earth. In Figure 2.8c, the negative rod is taken away. The negative charges are unable to return from earth, so the sphere is left positively charged. These charges now distribute themselves evenly.

Activity 2.7: Charging by induction

Tear a sheet of newspaper into small pieces, approximately half a centimetre in diameter. Place the bits under a glass plate that is supported by two books (see Figure 2.10).

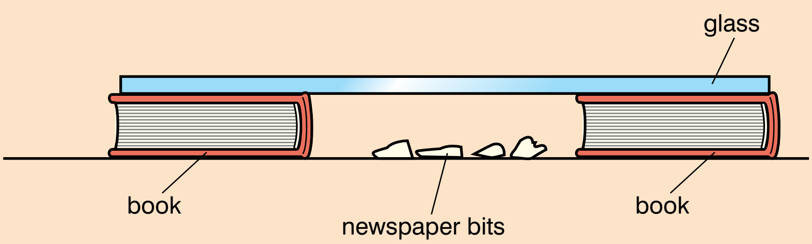


Figure 2.10 *Charging by induction.*

Rub the glass vigorously with a piece of silk and notice how the paper jumps up to the glass. In a small group, discuss what you see. Answer these questions:

- What attracts the paper to the glass?
- Why does it not stay attached to the glass?

Activity 2.6: Charging an electroscope by induction

From the information given about charging by induction, devise and carry out an experiment to charge an electroscope by induction. Draw diagrams to explain what happens.

DID YOU KNOW?

An electroscope is used to detect the presence and magnitude of electric charge. The British physician William Gilbert invented the first electroscope, which was called the *versorium*, around 1600. This was the first electrical measuring instrument.

Figure 2.9 *William Gilbert.*

KEY WORDS

conduction *the flow of charge from a charged object to an uncharged (neutral) object*

induction *a redistribution of electrical charge in an object, caused by the influence of nearby charges*

Figure 2.12 *Dr Robert Jemison Van de Graaf and Dr Karl Taylor Compton with a generator.*

KEY WORDS

conductor *a material that contains movable electric charges*

DID YOU KNOW?

The American physicist Robert J. Van de Graaf developed the **Van de Graaf generator** at Princeton University. The first model was demonstrated in October 1929 and used a silk ribbon bought at a local shop as the belt to carry the charge. Accelerating electrons are now used to sterilise food and process materials.

Activity 2.8: Using a Van de Graaf generator to make a fluorescent tube glow

Darken the room. Turn on the generator, and slowly move a fluorescent tube towards the generator. Why does the area between the tube, your hand and the generator glow?

Distribution of charges on the surfaces of conductors

Charges on the surface of a **conductor** are able to flow and spread themselves out evenly. The charges on the surface are all the same type and so they repel each other and move as far apart as possible. The result is that they move outwards until they reach the surface. They can get no further than that unless the insulation of the air breaks down to enable them to escape.

The charges on a sphere will distribute themselves evenly over its surface (see Figure 2.11a). If, however, the body has an irregular curvature (see 2.11b) the charge tends to cluster round the sharply curved parts, with the areas of gentle curvature storing far less of the charge. If the metal body has an actual point on it, the accumulation of charge there will be enormous, and the insulation of the air is more likely to break down there so that all the charge is able to escape from the surface. We refer to this is as the action of points.

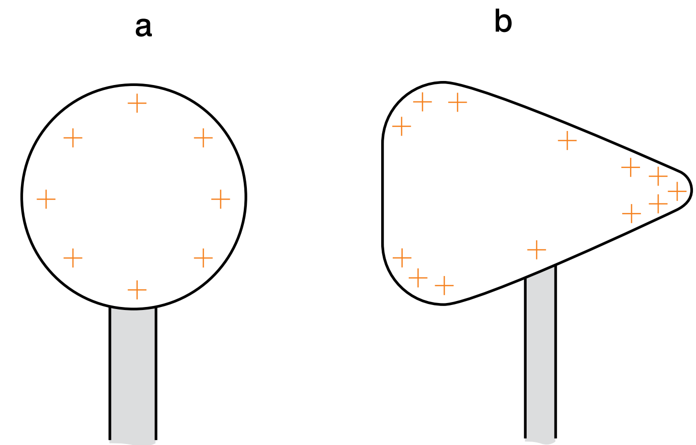


Figure 2.11 *Distribution of charges on conductors.*

As all the charge sits on the outside, it makes no difference whether the conductor, or metal object, is solid or hollow.

Thunderstorms

Scientists believe that thunderstorms in Ethiopia may sometimes trigger hurricanes as far away as across the Atlantic Ocean.

Thunderclouds form as a result of convection currents (warm air rises and cooler air goes down). Air contains water vapour which we often see as clouds.

As the air cools down, the water vapour in the air condenses. The tiny drops of water that condense experience very rough conditions and, in the process, tend to become charged. The mechanism by which this happens is more than just charging by friction, and we are still not entirely sure of all the details.

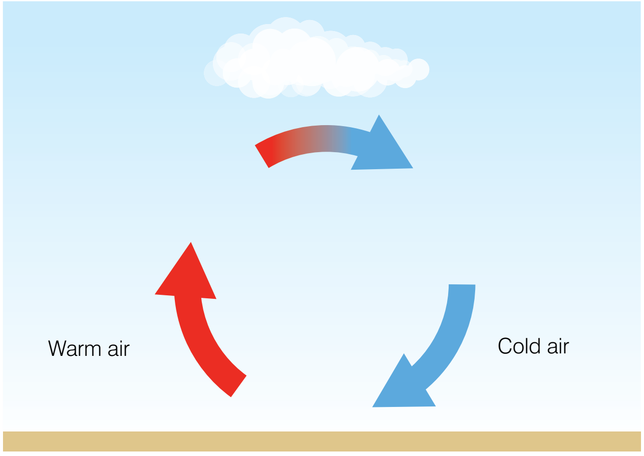


Figure 2.13 *A convection current.*

How lightning is formed

Charged water drops of one sign (positive or negative) may tend to collect in one part of the cloud. The result is lightning: sparks that jump often from one cloud region to another, but sometimes from the cloud to earth.

When a thundercloud passes overhead, if the base of the cloud is positively charged, then it will attract negative electrons in the earth underneath it (see Figure 2.15). Note that the charge carried by the base of the cloud may be negative or positive but the figure assumes that it is positive.

Should a spark jump to earth it will tend to jump the shortest air gap, such as to trees and tall buildings.

Lightning rods

To help to discharge the cloud safely, a lightning conductor is often fitted to the top of a building. This is a metal bar, pointed at the top, the other end of which is buried firmly in the ground. The idea is that it should get struck before the building and conduct the surge of charge harmlessly to the earth.

Figure 2.16 *A lightning conductor will conduct electric charge from lightning safely away from a building to the earth.*

DID YOU KNOW?

Fluorescent lamps work because there is a flow of electrical current through a gas (usually argon) which has been charged.

KEY WORDS

Van de Graaf generator *a hollow sphere on which an electrostatic charge is accumulated using a moving belt*

Figure 2.14 *Lightning.*

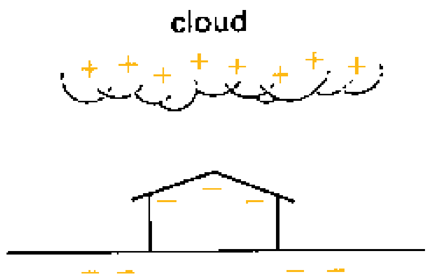


Figure 2.15 *How lightning occurs.*

Activity 2.9: How might tall buildings be protected from damage by lightning?

In a small group, discuss how tall buildings might be protected from damage by lightning.

Activity 2.10: Why is the inside of a car safe in a storm?

Discuss with a partner why the inside of a car is a safe place to shelter during a storm.

Activity 2.11: Design a safety poster

Use the information above to design a poster to show younger children how to stay safe during a thunderstorm.

Safety in storms

It is very rare that people are struck by lightning, and certainly you will not be struck while you are inside a car or an aeroplane – the metal shell around you would divert charge away. You are advised not to shelter under trees because they are particularly likely to get struck, and the spark could jump from the lower branches to you.

Applications of electrostatics

Paint sprays

Electrostatic spray painting is often used to paint car body panels and bike frames. An electrostatic spray painting device sprays charged particles of paint through the air onto a surface. These devices are usually used for covering large surfaces with an even coating of liquid. They can be either automated or hand-held and have interchangeable heads to allow for different spray patterns.

There are three main technologies for charging the fluid (liquid or powders):

- *Direct charging* – an electrode (a charged metal plate) is placed in the paint supply reservoir or in the paint supply conduit.
- *Tribo charging* – this uses the friction of the fluid, which is forced through the barrel of the paint gun. It rubs against the side of the barrel and, in the process, builds up an electrostatic charge.
- *Post-atomisation charging* – the charged fluid comes into contact with an electrostatic field (see page 54) after leaving the spray head. The electrostatic field may be created by electrostatic induction, or by one or more electrodes (electrode ring, mesh or grid).

The charged paint particles repel each other, and so they spread themselves evenly as they leave the spray head. The object being painted is charged in the opposite way (for example, if the paint is positively charged, the object is negatively charged), or grounded. The paint is then attracted to the object, which gives a more even coat than the alternative technique, wet spray painting, and increases the percentage of paint that actually sticks to the object. This method also means that paint covers hard to reach areas. The object is then baked so that the paint sticks properly when the paint powder particles turn into a type of plastic.

Figure 2.17 *An electrostatic spray gun in use.*

Chimney filters

Smoke such as that from a large chimney contains many tiny solid particles, mainly soot. To help to keep the air clean, before reaching the chimney this smoke may be passed between two large upright charged metal plates. In a similar way to how a plastic comb attracts small pieces of paper, the solid specks in the smoke are trapped by sticking to one or other of the plates. From time to time the plates are shaken or scraped to remove the thick layer of dust that has formed on them.

Photocopiers

1. The surface of a cylindrical drum is charged using electrostatics. The drum has a coating of a material that conducts electricity when light shines on it.

2. A bright lamp shines light on the original document, and the white areas of the original document reflect the light onto the surface of the charged drum. The areas of the drum that are exposed to light (those areas that correspond to white areas of the original document) conduct electricity so they discharge. The areas of the drum that are not exposed to light (those areas that correspond to black portions of the original document) remain negatively charged. The result is a stored electrical image on the surface of the drum. (In digital machines, the original document is scanned and digitised and a laser is employed to discharge the drum in a similar fashion.)

3. The toner is positively charged. When it is applied to the drum to produce the image, it is attracted and sticks to the areas that are negatively charged (black areas), just as the balloon in Activity 2.4 stuck to the wall or door.

4. The toner image on the surface of the drum is transferred from the drum onto a piece of paper, which is charged more negatively than the drum.

5. The toner is melted and bonded to the paper by heat and pressure rollers.

6. The drum is wiped clean with a rubber blade and light is used to remove the charge.

Activity 2.12: Design a model chimney filter

In a small group, design a model of chimney filter.

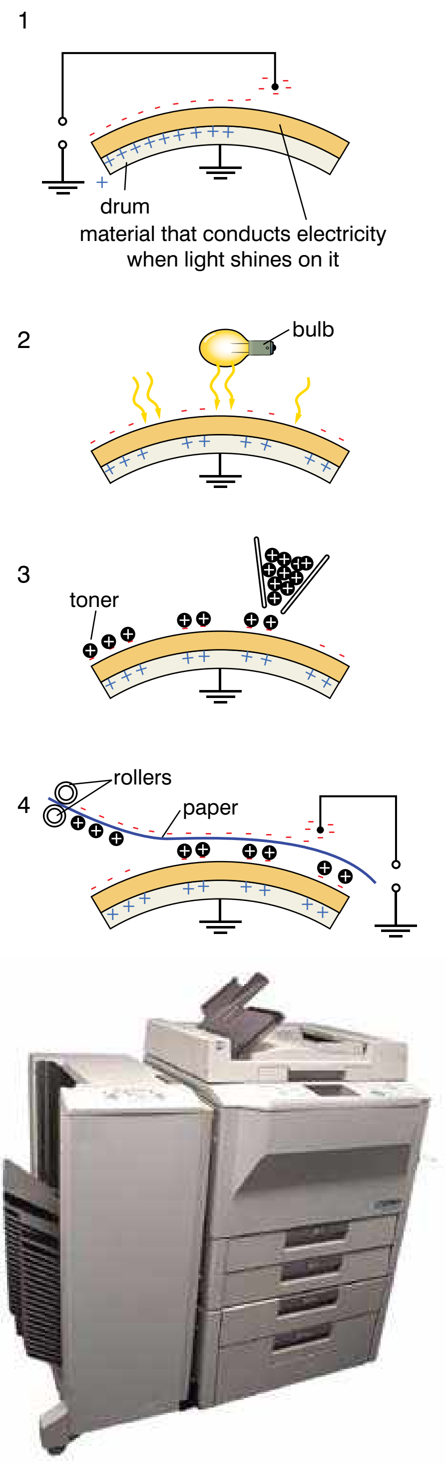


Figure 2.18 How a photocopier works.

When does electrostatic charge cause problems?

Electrostatic charge can be a problem or hazard when there is a sudden discharge. Electrostatic discharge gives a sudden current of electricity between two objects. On page 56, we found out that lightning (which is a form of electrostatic discharge) can be dangerous, especially if you are outside in a thunderstorm. If there is an electrostatic discharge in an area containing liquids or materials which catch fire easily (flammable materials), then the discharge may cause a fire. However, perhaps the main area where electrostatic discharge causes problems is in electronics, where circuits may be damaged by unwanted flows of charge. To protect circuits against this, antistatic devices have been developed. These include antistatic bags, clothing, mats and wrist straps.

Why is static electricity more apparent when the air is dry?

You notice static electricity much more when the air is dry because dry air is a relatively good electrical insulator, so if something is charged the charge tends to stay. When there is moisture in the air, the charged water molecules can remove charge quickly from a charged object.

Activity 2.13: Mind mapping uses and hazards of electrostatics

In a small group, discuss what you have learnt about the uses and hazards of electrostatics. You may like to do some research to find out how electrostatics is used in manufacturing, or more about antistatic devices. Make a summary of your discussion and research using mind maps such as the one in Figure 2.19.

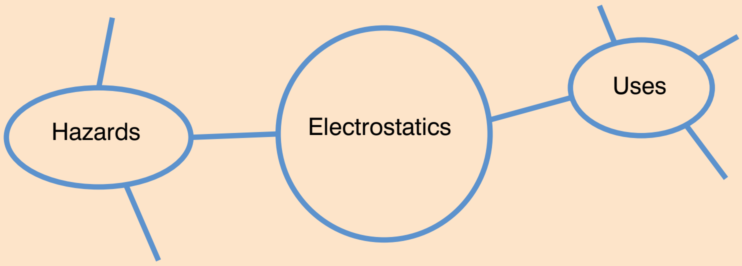


Figure 2.19 *Mind map to show uses and hazards of electrostatics.*

Prepare a presentation for the rest of your class.

Summary

- Charge cannot be created or destroyed, it can only be transferred from one object to another. This is the law of conservation of charge.
- It is possible to charge objects by rubbing, conduction and induction.
- An electroscope can be charged by conduction and induction.
- The distribution of charge on a conductor with a constant surface such as a sphere is uniform (the charges spread out evenly). However, if there is a point on the surface, the charge will form clusters around the point.

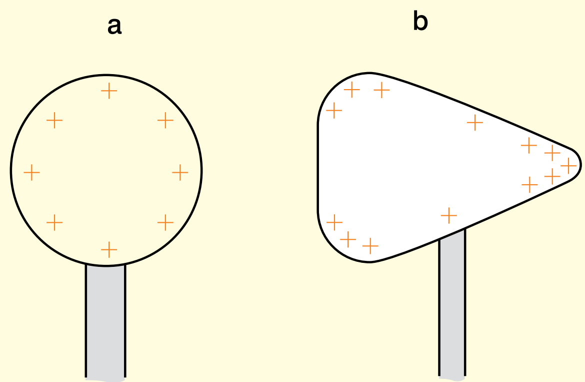


Figure 2.20 *Distribution of charge on conductors of various shapes.*

- Lightning is formed when charged water drops of one sign (positive or negative) collect in one part of the cloud. There is electrostatic discharge (in the form of sparks) which jump often from one cloud region to another, but sometimes from the cloud to earth.
- A lightning rod may be used to protect tall buildings from lightning. This is a metal rod with a point at the top. It attracts the charge from the lightning and conducts it safely to earth rather than it going through the building and causing damage.
- Electrostatics is used in equipment such as paint sprayers and photocopiers. However, it can be a problem in some situations, such as where there are flammable materials, which may catch fire if there is an electrostatic spark, or in sensitive electronic devices. There are various antistatic devices available, which can be used in different situations.

Review questions

1. Why do charged objects soon lose their charge on a damp day?
2. All atoms that make up matter contain tiny charges, both positive and negative. How is it that most objects we meet in daily life appear to be uncharged?
3. Copy Figure 2.21 and show on it how the charges in the uncharged object will distribute themselves in the presence of the positively charged rod.

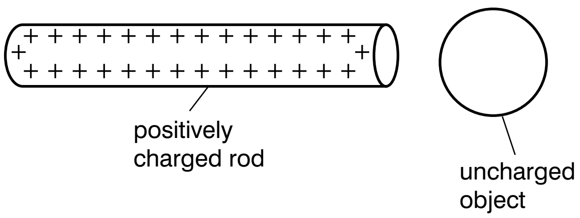


Figure 2.21

4. Why will a plastic comb that has been rubbed attract small pieces of paper?
5. a) Why do plastic objects frequently become covered with dust?
   b) Why is it not necessarily a good idea to rub such objects with a duster?
6. When car manufacturers apply paint to new cars they usually spray it in very fine droplets that are deliberately charged. Why do they do this?
7. a) An inflated toy balloon can be rubbed on your clothing then stuck on the ceiling. Explain why.
   b) Suggest why the balloon eventually drops off the ceiling.

8. Why are metal objects designed to store charge (e.g. the dome of a Van de Graaf generator) made from large hollow metal spheres?
9. Explain the purpose of a lightning conductor, and draw a diagram to show how it does its job.
10. a) List some uses of electrostatics
    b) List some hazards of electrostatics.
11. List some antistatic devices.

2.2 Electric forces and fields

By the end of this section you should be able to:

- Define an electric field.
- Represent diagrammatically the electric field lines around and between two points.
- Distinguish between the electric field inside, outside and between surfaces of a spherical metallic conductor.
- State Coulomb's law.
- Compare Coulomb's law and Newton's law of universal gravitation.
- Calculate the force acting on a charge due to two other charges placed on the same plane (line of action).
- Calculate the force between three charges placed in a line.
- Calculate the electric field strength at a point due to charges placed in a line and at right angles.

KEY WORDS

electric field *a region where an electric charge will experience a force which is due to the presence of other electric charges*

Also see…

The material in this section is related to Unit 4, Section 4.4.

What is an electric field?

Imagine you are a positively charged particle. Whenever you are anywhere near other charges you will find yourself in a world of pushes or pulls. We say you are in an **electric field**.

An electric field is a region where an electric charge will experience a force that is due to the presence of other electric charges.

Clearly a charge will be surrounded by an electric field, because everywhere around it other charges will either be attracted to it or repelled from it, depending on their signs.

Plotting electric field lines

Electric fields can be plotted by indicating the direction of the force on a positive charge, if one happened to be there. Figure 2.22a shows the electric field around a positive charge, and 2.22b shows that round a negative charge.

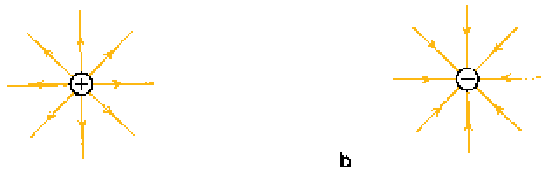


Figure 2.22 Electric fields around a positive charge, and a negative charge.

Notice two things. First, a negative charge, such as an electron, would move in the opposite way to the direction shown by the arrows. Second, the field is strongest where the lines are closest together.

We can also plot the field around two charges. In Figure 2.23a you can see the field around a pair of opposite charges, while 2.23b shows the completely different situation when both charges are the same sign.

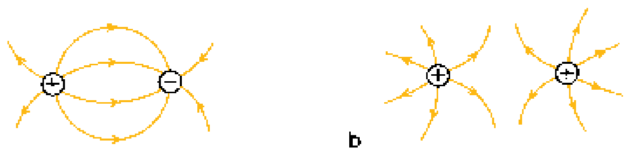


Figure 2.23 Electric fields around a pair of opposite charges, and a pair of charges with the same sign.

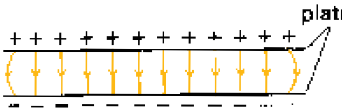


Figure 2.24 The uniform field between a parallel pair of charged plates.

Any charge between a pair of charged plates would find itself in a field that was both strong and uniform, as shown in Figure 2.24.

Activity 2.14: Explaining field plots

With a partner, see if you can explain the plots of electric fields shown on this page. Do it by imagining that, in turn, you or your partner are a tiny positively charged particle placed at different points in the field. Think about both the direction of the force you feel and how strong it is.

Also see…

Compare the material on this page to that in Unit 4, Section 4.2.

Electric field inside and outside a spherical metallic conductor

Imagine a spherical, hollow conductor such as the one shown in Figure 2.25.

Activity 2.15: Explaining field patterns

Work with a partner to see whether you can explain the field pattern in Figure 2.25 before reading on!

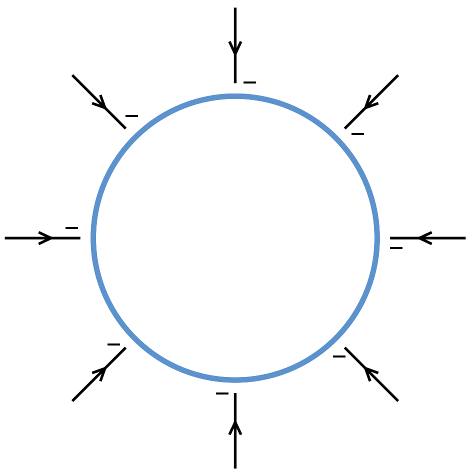


Figure 2.25 Electric field inside and outside a charged hollow spherical metallic conductor.

Inside the conductor there is no charge because the charge has collected on the outer surface. This means that the inside of such a conductor can be used to shield equipment from external electric fields. The outer surface of the conductor will have an electric field that is at right angles to (perpendicular to) the surface.

Electric field strength

An electric charge that finds itself in an electric field will experience a force. There are two factors that determine how large a force F this will be:

- The size of the charge Q, measured in **coulombs**.
- The strength (intensity) of the electric field, which corresponds to the closeness of the field lines.

We define the **electric field strength** (electric field intensity) E by:

$$E = \frac{F}{Q}$$

The units of E will be newtons of force per coulomb of charge – that is, N/C; E is a vector – it has direction as well as magnitude.

Worked example 2.1

A charge of 5 C is in an electric field of 3 N C^{-1}. Find the force acting on the charge.

E (N/C)	*F* (N)	*Q* (C)
3	?	5

Rearranging $E = \frac{F}{Q}$ gives $F = QE$. Putting in the values, the force acting on the charge will be

5 C × 3 N/C = 15 N.

(Notice that the units must work out, as well as the numbers.) If it is a positive charge, the force will act in the direction of the arrow on the field lines. The force on a negative charge will be in the opposite direction.

Worked example 2.2

A negative charge of 8 C is in an electric field of 4 N/C. Find the force acting on the charge.
In which direction will it act?

E (N/C)	*F* (N)	*Q* (C)
4	?	–8

Rearranging $E = \frac{F}{Q}$ gives $F = QE$. Putting in the values, the force acting on the charge will be

8 C × 4 N/C = 32 N.

Since it is a negative charge, the force will act in the opposite direction to the direction of the arrow on the field lines.

DID YOU KNOW?

Charles-Augustin de Coulomb, a French physicist, is best known for developing the definition of the **electrostatic force** of attraction and repulsion. This is summarised in Coulomb's law (see page 64). The coulomb is the SI unit of charge and was named after him. During the French Revolution he took part in the determination of weights and measures. He was one of the first members of the French National Institute. As well as his contribution to the understanding of electric fields, Coulomb has contributed to the design of retaining walls!

Figure 2.26 *Charles-Augustin de Coulomb, 1736–1806.*

KEY WORDS

electrostatic force *the force exerted by stationary objects bearing electric charge on other stationary objects bearing electric charge*

Worked example 2.3

Find the electric field strength when a positive charge of 7 C is acted on by a force of 28 N. In which direction will the force act?

E (N/C)	*F* (N)	*Q* (C)
?	28	7

Put the given values into $E = \frac{F}{Q}$ so $E = \frac{28}{7}$.

The electric field strength will be 4 N/C.

Since it is a positive charge, the force will act in the direction of the arrow on the field lines.

Worked example 2.4

Find the value, and type, of charge that is acted on by a force of 40 N acting in the opposite direction to the arrow on the field lines in an electric field of strength 10 N C^{-1}.

E (N/C)	*F* (N)	*Q* (C)
?	28	7

Rearrange $E = \frac{F}{Q}$ to $Q = \frac{F}{E}$. Put the given values into $Q = \frac{F}{E}$ so Q = 40/10. The charge will be 4 C.

Since the force acts in the opposite direction to the the arrow on the field lines, it is a negative charge.

Reminder: rearranging equations

Remember that you can use a triangle like the one in Figure 2.27 to rearrange the equation connecting three quantities.

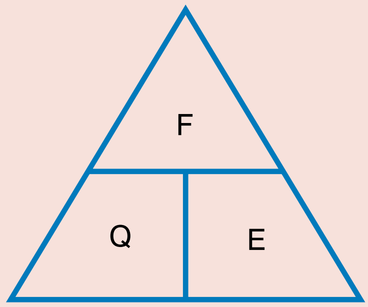


Figure 2.27 *To find, for example, Q, cover up Q in the triangle and then you see that $\frac{F}{E}$ is left. This is the formula you need!*

KEY WORDS

coulomb *unit of electric charge*

electric field strength *the electrostatic force acting on a small positive test charge placed at that point*

Force between charges

We know that two positive electric charges Q_1 and Q_2 a distance *r* apart will repel each other. This is because each one lies in the field created by the other (see Figure 2.28).

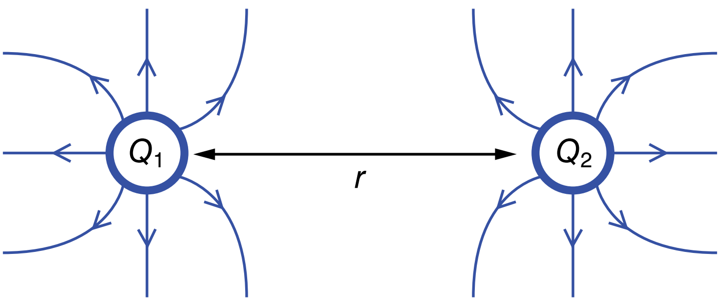


Figure 2.28 *Q_1 lies in the field of Q_2 and, since both charges are positive, they will repel each other.*

It has been found by carrying out experiments that:

- The force *F* between the charges increases if the value of either of the charges is increased.
- The force gets less as the distance between the charges gets larger, and does so by an inverse relationship. The force decreases in relation to the inverse of the distance squared, $\frac{1}{r^2}$.

Activity 2.16: Comparing Coulomb's law and Newton's law of universal gravitation

With a partner, recall Newton's law of universal gravitation. (Try to do this without looking back to page 41!) Write the two formulae side by side. Compare them. Make a list of things that are similar and things that are different.

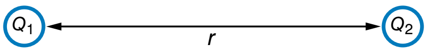


Figure 2.29 *Coulomb's law.*

This means that if the separation doubles, the force reduces to one quarter; if it increases three times, the force reduces to one ninth, and so on.

Coulomb's law states that the force between two point charges is directly proportional to the product of the two charges and inversely proportional to the square of the distance between the two.

We can express this in a formula as $F = \frac{k \times Q_1 Q_2}{r^2}$,

where k is a constant. The newton, the coulomb and the metre are already fixed; that constant has to be obtained by measurement. You might expect the constant to be given a symbol such as 'E', perhaps, but instead it is written as $\frac{1}{4\pi\varepsilon_0}$. Do not be put off by this: it is still a constant, and nothing more complicated than that.

This gives the final expression as:

$$F = \frac{1}{4\pi\varepsilon_0} \times \frac{Q_1 Q_2}{r^2}$$

Check that you understand why the units for the constant $\frac{1}{4\pi\varepsilon_0}$ will be N m²/C². The value of this constant depends on the material in which the charges are placed.

Also see...

The material on this page may be compared to that in Unit 1, Section 1.4.

Worked example 2.5

The value of $\frac{1}{4\pi\varepsilon_0}$ in a vacuum is 9.0×10^9 N m² C⁻². What is the force between two charges, 5 C and 7 C, placed 1 metre apart in a vacuum?

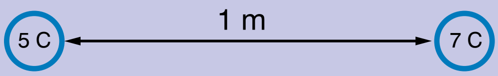


Figure 2.30 *A 5 C charge and a 7 C charge placed 1 metre apart in a vacuum.*

Use $F = \frac{1}{4\pi\varepsilon_0} \times \frac{Q_1 Q_2}{r^2}$

F (N)	**Q_1 (C)**	**Q_2 (C)**	**r (m)**
?	5	7	1

Substitute the given values: Q_1 = 5 C and Q_2 = 7 C, the constant value is 9.0×10^9 N m²/C and, since $r = 1$, $r^2 = 1$.

So $F = 5 \times 7 \times (9.0 \times 10^9 \text{ N m}^2/\text{C}^2) = 3.15 \times 10^{11}$ N

KEY WORDS

Coulomb's law *the force between two point charges is directly proportional to the product of the two charges and inversely proportional to the square of the distance between the two*

Activity 2.17: Explaining how results of calculations demonstrate the inverse square relationship

Explain to a partner how the results of the two worked examples on this page demonstrate the inverse square relationship suggested in Coulomb's law.

Try to demonstrate the inverse square relationship suggested in Newton's law of universal gravitation using similar examples. (Take the value of $G = 6.67 \times 10^{-11}$ N m²/kg².)

Worked example 2.6

The value of $\frac{1}{4\pi\varepsilon_0}$ in a vacuum is 9.0×10^9 N m^2 C^{-2}. What is the force between two charges, 5 C and 7 C, placed 2 metres apart in a vacuum?

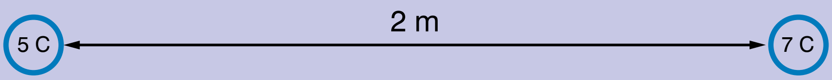


Figure 2.31 *A 5 C charge and a 7 C charge placed 2 metres apart in a vacuum.*

Use $F = \frac{1}{4\pi\varepsilon_0} \times \frac{Q_1Q_2}{r^2}$

Substitute the given values: $Q_1 = 5$ C and $Q_2 = 7$ C, the constant value is 9.0×10^9 N m^2 C^{-2} and, since $r = 2$, $r^2 = 4$.

So $F = \frac{[5 \times 7 \times (9.0 \times 10^9 \text{ N m}^2/\text{C}^2)]}{4} = 7.875 \times 10^{10}$ N

Coulomb's law when there are more than two charges

Electric force has both size and direction. The two electric forces in Figure 2.32 are the same size but opposite in direction. Charges of the same sign exert repulsive forces on one another, while charges of opposite sign attract, so these two charges would repel each other.

Figure 2.32 *Forces of the same size but in opposite directions.*

When more than one charge exerts a force on another charge, the total force on that charge is the sum of the individual forces, taking into account both their sizes and direction.

Worked example 2.7

Three charges are arranged in a line with 2.5 cm between them, as shown in Figure 2.33.

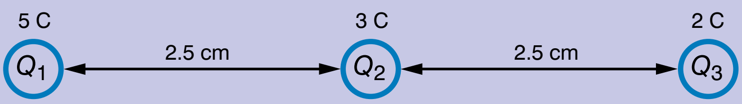


Figure 2.33 *Three charges in a line at 2.5 cm apart.*

Q_1 is 5 C, Q_2 is 3 C and Q_3 is 2 C, $k = 8.99 \times 10^9$ N m^2/C^2.

What is the force exerted on Q_2 by the other two charges?

First draw a good diagram showing the forces acting on the charge. The diagram should also show the directions of the forces.

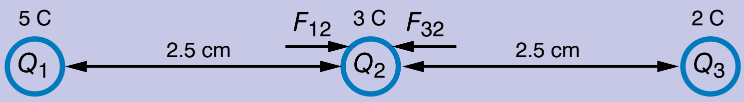


Figure 2.34 *Forces acting on a charge with their direction shown.*

Consider the forces exerted on Q_2 by the other two:

from Q_1: $F_{12} = \dfrac{kQ_1Q_2}{r^2} = \dfrac{8.99 \times 10^9 \times 5 \times 3}{(0.025)^2} = 2.1576 \times 10^{14}$

from Q_3: $F_{32} = \dfrac{kQ_3Q_2}{r^2} = \dfrac{8.99 \times 10^9 \times 2 \times 3}{(0.025)^2} = 8.6304 \times 10^{13}$

The total force on Q_2 is therefore $2.1576 \times 10^{14} + 8.6304 \times 10^{13} = 3 \times 10^{14}$ (to 1 significant figure)

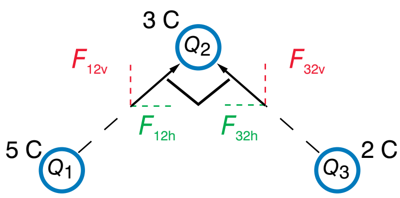


Figure 2.35 *Finding the forces between charges.*

Note that if Q_1 and Q_3 are at right angles to Q_2 as shown in Figure 2.35, you find the forces between the charges in the same way as in the example above but the lines of action of the forces do not coincide so you have to add the horizontal components of the total force and the vertical components of the total force. Note: F_{12v} is the vertical component of the force between Q_1 and Q_2.

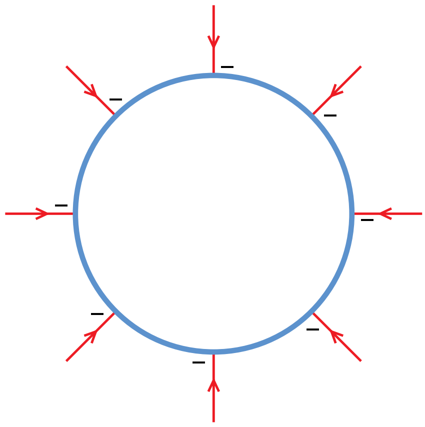


Figure 2.36 *An electric field on the inside and outside surfaces.*

Summary

- An electric field is a region where an electric charge will experience a force because other charges are present.
- The electric field on the inside and outside surfaces of a spherical metallic conductor is as shown in Figure 2.36.
- Coulomb's law states that the force between two point charges is directly proportional to the product of the two charges and inversely proportional to the square of the distance between them.
- We can express this in a formula as $F = \dfrac{k \times Q_1Q_2}{r^2}$, where k is a constant.
- Coulomb's law and Newton's law of universal gravitation are both inverse square relationships.
- You can calculate the force between three charges placed in a line.

Review questions

1. Explain, as if to a friend, the meaning of the term 'electric field'.
2. Draw the pattern of the electric field around:
   a) a positive charge
   b) a negative charge.

   Explain what the arrows mean and explain at which points the strength of the field is greatest.
3. Repeat question 2 for the electric field:
   a) around a positive and a negative charge
   b) around two negative charges
   c) between two oppositely charged flat metal plates.
4. What is the name of the unit we use to measure electric charge? Give its correct abbreviation.
5. How big a force will a charge of +2 C experience if placed in an electric field of strength 18 N/C?
6. In question 5, what can you say about the force if the charge was one of −2 C instead?
7. A charge is placed in an electric field of 48 N/C and it experiences a force of 12 N. How great must the charge be?
8. Two point charges, one of 3.0 C and the other of 2.0 C, are placed 5.0 m apart. Calculate the force which acts between them. (Take $\frac{1}{4\pi\varepsilon_0}$ to be 9.0×10^9 N m^2 /C^2.)
9. The answer to question 8 is an *enormous force* – a coulomb is a huge unit for static charge, and we do not normally encounter point charges of anything like that size. Repeat the calculation but this time take the charges as 3.0 pC (p = pico- = $\times 10^{-12}$) and 2.0 μC.

2.3 Electric potential

By the end of this section you should be able to:

- Define electric potential and its SI unit.
- Distinguish between absolute potential and potential difference.
- Show that 1 N/C = 1 V/m.
- Explain equipotential lines and surfaces.
- Draw equipotential lines and surfaces in an electric field.
- Define the term electric potential energy.

DID YOU KNOW?

Alessandro Volta (1745–1827) was an Italian physicist known especially for the development of the first electric cell. Volta made contributions for chemistry: in 1776–77 he discovered methane by collecting the gas from marshes. He devised experiments such as the ignition of methane by an electric spark in a closed vessel.

Volta began to study, around 1791, the 'animal electricity' noted by Luigi Galvani when two different metals were connected in series with a frog's leg and to one another. He replaced the frog's leg by brine-soaked paper, and detected the flow of electricity.

In 1800, he invented the voltaic pile, an early electric battery, which produced a steady electric current.

Figure 2.37 *Alessandro Volta.*

Electric potential and its SI unit

Electric potential at a point in space is potential energy divided by charge that is associated with a static (one that does not vary with time) electric field. Its SI unit is the volt, in honour of Alessandro Volta.

You know from earlier in this unit that objects may possess electric charge. An electric field exerts a force on charged objects, accelerating them in the direction of the force, in either the same or the opposite direction of the electric field. If the charged object has a positive charge, the force and acceleration will be in the direction of the field. This force has the same direction as the electric field vector, and its magnitude is given by the size of the charge multiplied with the magnitude of the electric field.

The electric potential created by a point charge q, at a distance r from the charge (relative to the potential at infinity), can be shown to be

$$V = \frac{1}{4\pi\varepsilon_0} \times \frac{q}{r}$$

where ε_0 is the electric constant.

Absolute potential and potential difference

To define absolute potential you need a reference point (for example, infinity). Then you can say that:

electric potential at a point = work done per unit charge in bringing a small object from infinity to a point in an electric field

To define electric potential difference between two points P and Q, you need to assume the absolute electric potential at P is V_P and the absolute electric potential at Q is V_Q. Then the electric potential difference is:

$$V_P - V_Q$$

How to show that that two quantities are equivalent

You can show that two quantities are equivalent by using their units.

From the definition of V we can find its units in terms of N, m and C.

$$V = \frac{1}{4\pi\varepsilon_0} \times \frac{q}{r}$$

We know that the units for the constant term are N m^2/C^{-2}.

The units for $\frac{q}{r}$ are C/m. So the units for V are $\frac{\text{N m}^2\text{ C}}{\text{C}^2\text{ m}}$.

This can be simplified, using the laws of indices, to N m/C.

If we divide this by m, to give V/m, we get N/C. So 1 V/m = 1 N/C.

Equipotential lines and surfaces

The term 'equi' means equal. Equipotential therefore means 'equal potential'. Equipotential lines are lines showing where the electric potential is the same value. Equipotential surfaces are surfaces that have the same electric potential.

Equipotential lines and surfaces in an electric field

You have already seen a diagram like that in Figure 2.38. The electric field strength along the lines marked in blue is the same at all points along the line. These are therefore lines of equipotential.

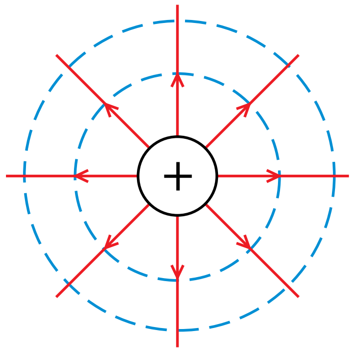


Figure 2.38 *Equipotential lines are shown in blue.*

Electric potential energy

Potential energy is energy that a body possesses because of its position. If you have a spring and you compress it, you have to do work to make the compression. The work you do is transferred to the spring as potential energy. In a similar way, when a positive charge (say) is moved towards another positive charge, they will repel each other, so work has to be done to push them towards each other. When there is no longer any work being done, the charges will be able to repel each other (the natural state). Electric potential energy is the energy a charge possesses because it is in the region of other charges.

Note that you can convert one form of energy (e.g. electric potential energy) to another form (e.g. kinetic energy, which is energy relating to movement) but energy cannot be created or destroyed. Energy is conserved in a system.

Summary

- Electric potential at a point in space is potential energy divided by charge that is associated with a static (one that does not vary with time) electric field. Its SI unit is the volt, in honor of Alessandro Volta.
- To define absolute potential you need a reference point (for example, infinity). Then you can say that:
  electric potential at a point = work done per unit charge in bringing a small object from infinity to a point in an electric field
- To define electric potential difference between two points P and Q, you need to assume the absolute electric potential at P is V_P and the absolute electric potential at Q is V_Q. Then the electric potential difference is:

  $V_P - V_Q$
- You can show that 1 N/C = 1 V/m.
- Equipotential lines are lines showing where the electric potential is the same value. Equipotential surfaces are surfaces that have the same electric potential.
- You can draw equipotential lines and surfaces in an electric field.
- Electric potential energy is the energy a charge possesses because it is in the region of other charges.

Activity 2.18: Drawing equipotentials

Using Figures 2.22b, 2.23a and 2.23b on page 61, draw diagrams showing equipotential lines around
a) a negative charge,
b) two opposite charges,
c) two positive charges.

KEY WORDS

electric potential *potential energy divided by charge that is associated with a static electric field. Its SI unit is the Volt*

Review questions

1. Define electric potential and its SI unit.
2. Distinguish between absolute potential and potential difference.
3. Show that 1 N/C = 1 V/m.
4. Explain equipotential lines and surfaces.
5. Draw equipotential lines between two negative charges in an electric field.
6. Define the term electric potential energy.

2.4 Capacitors and capacitances

By the end of this section you should be able to:

- Describe the structure of a simple capacitor.
- Define the term capacitance and its SI unit.
- Apply the definition of capacitance to solve numerical problems.
- Use a circuit symbol to represent a capacitor.
- Explain the charging and discharging of a capacitor.
- Define the term dielectric and explain what is meant by a dielectric material.
- Identify combination of capacitors in series, parallel and series–parallel.
- Explain the effect of inserting a dielectric in the gap between the plates of a parallel plate capacitor.
- Derive an expression for the effective capacitance of capacitors connected in series and parallel.
- Draw an electric circuit diagram for a simple capacitor, series and parallel connections of two or more capacitors using symbols.
- Solve problems on combination of capacitors.
- Define parallel plate capacitor.
- Describe the factors that affect the capacitance of a parallel plate capacitor.
- Calculate the capacitance of a parallel plate capacitor.
- Find an expression for the electric potential energy stored in a capacitor.
- Calculate the energy stored in a capacitor using one of three possible formulae.
- State some uses of capacitors in everyday life.

The structure of a simple capacitor

A capacitor is a small device designed to store more charge at a lower potential. The commonest way of doing this is to use two parallel plates, a tiny distance apart and separated by an insulator (which may be air or may be a dielectric material – see page 73).

Figure 2.39 *A selection of capacitors.*

Capacitance and its SI unit

If we place some charge on an insulated metal sphere, the sphere's voltage will rise; put some more charge on, and its voltage will rise further.

Presumably a larger sphere will hold more charge before its potential has risen by 1 V. We measure this by using a quantity called capacitance, C. If a charge Q results in a rise of V in the potential, we define the capacitance of the sphere by:

$$C = \frac{Q}{V}$$

This is the charge needed for each volt rise in the sphere potential. The units will be coulombs per volt, C/V, which we call farads, F.

The farad is a very large unit of capacitance – it is rare for an extra coulomb of charge to result in a rise in potential of no more than 1 V. We often use the microfarad (μF, $\times 10^{-6}$ F) or even the picofarad (pF, $\times 10^{-12}$ F).

Worked example 2.8

The metal dome at the top of a Van de Graaf generator will have a capacitance of around 10 pF (1×10^{-11} F). The insulation of air will typically break down when subjected an electric field of around 3000 V/mm. Suppose that the charged dome will produce sparks to earth across an air gap of about 33 mm. That suggests that the potential of the sphere is something like 3000 V/mm $\times$ 33 mm, which gives 100 000 V.

C (F)	Q (C)	V (V)
1×10^{-11}	?	1×10^{5}

The charge held on the dome can be found using

$C = \frac{Q}{V}$.

Rearranging, $Q = CV$

$= 1 \times 10^{-11}$ F (C/V) $\times 1 \times 10^{5}$ V

$= 1 \times 10^{-6}$ C (1 μC)

Worked example 2.9

A 150 pF capacitor holds 6.0×10^{-9} C of charge. What will the p.d. be between its terminals?

C (F)	Q (C)	V (V)
150×10^{-12}	6×10^{-9}	?

$C = \frac{Q}{V}$

We want to find V. Rearrange to give

$V = \frac{Q}{C}$

Substitute values

$$V = \frac{6.0 \times 10^{-9}}{150 \times 10^{-12}} = 40 \text{ V}$$

Larger spheres

A larger sphere will have a bigger capacitance, so you will need to give it a greater amount of charge before its potential rises by 1 V. However, a sphere with a capacitance of 1 F would be vast – its radius would be around 2000 times that of the Earth.

Activity 2.19: Charging and discharging capacitors

Charge a large capacitor for a few minutes by connecting it to a battery, as shown in Figure 2.41.

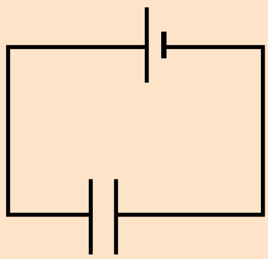


***Figure 2.41** Circuit showing capacitor connected to a battery.*

Then add a bulb to the circuit, as shown in Figure 2.24.

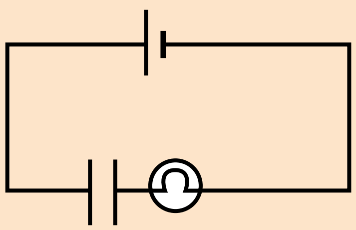


***Figure 2.42** Circuit with bulb.*

What happens? Try to draw a graph showing light intensity versus time. Discuss what was stored and what was drained before reading on.

Also see...

See page 94 for more on Ohm's law.

The circuit symbol for a capacitor

Figure 2.40 shows the usual circuit symbol for a capacitor.

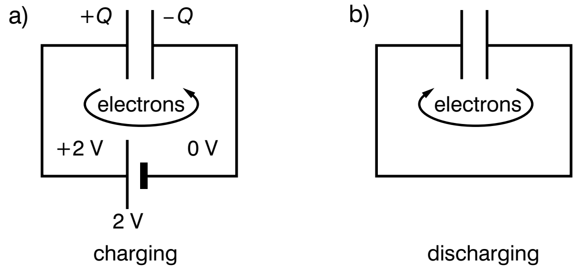


***Figure 2.40** Circuit symbol for a capacitor.*

In Figure 2.40a , the battery is moving electrons from the plate $+Q$, so it is left with a positive charge. The electrons move round the wires to the $-Q$ plate, which is left with a negative charge. When the potential difference between the plates is the same as that across the battery (2 V), no more charge can flow because the capacitor is balancing the battery.

How much charge is being stored on the capacitor? It might be argued that for example, with $+Q$ on one plate and $-Q$ on the other as shown above, the total amount stored is $2Q$. More plausibly, perhaps, somebody else might say $+Q$ and $-Q$ means a total charge of zero. However, neither is correct. Look at Figure 2.40b which shows the capacitor emptying. What happens is that 'Q' coulombs of electrons flow off one plate, pass round the circuit in the form of a momentary current and go on to the other plate. The charge stored is Q.

The formula for capacitance is still $C = \frac{Q}{V}$ but the terms now take on a slightly different meaning: Q is the charge held on one of the plates, and V is the potential difference between them.

Charging and discharging a capacitor

When a capacitor discharges through a resistor (or a bulb), the capacitor acts like a battery to drive a current through the resistor, but, unlike a battery, its voltage drops rapidly as its charge drains away. Think of a capacitor charged up to a voltage V_0 which is then emptied through a resistor R. When it is first connected, the discharging current is determined entirely by the resistor (since $I = \frac{V}{R}$ for it) – the higher the value of the resistor, the more slowly the capacitor will empty.

Figure 2.43 indicates how the charge remaining on the capacitor will decrease with time; the vertical axis actually shows the voltage across its plates, but this also provides a measure of the charge left.

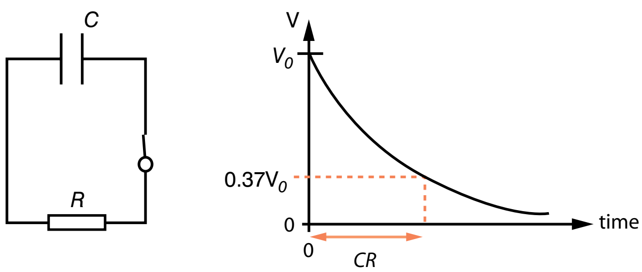


***Figure 2.43** Charge decreasing over time.*

As the capacitor empties, so the voltage across its plates drops. This decreasing voltage across the resistor *R* results in a reducing discharge current, so the capacitor empties ever more slowly. The resulting curve is what we call exponential decay.

We can work out how long it takes for the voltage to drop to 0.37 of its starting value (by which time the capacitor has almost two-thirds emptied). We do this by multiplying *C* by *R*. This gives what we call the time constant of the circuit.

Working out *CR* gives a time in seconds. This means that F Ω, farads multiplied by ohms, must be seconds. It seems unlikely, but it is certainly a good exercise to try. From the definition of capacitance, be sure you can see why the farad is coulombs per volt, C/V. The ohm is derived from $R = \frac{V}{I}$, and so is volts per ampere, V/A.

Multiply them together and the unit reduces to C/A. Realise that the coulomb is the ampere second, A s, and you are there.

When a capacitor is charged through a resistor, we have the opposite to the discharge situation. Here the capacitor charges rapidly at the start, but then continues at an ever-declining rate. It is a kind of 'upside down' exponential curve. When the voltage between the plates of the capacitor equals that of the battery, the charging ceases.

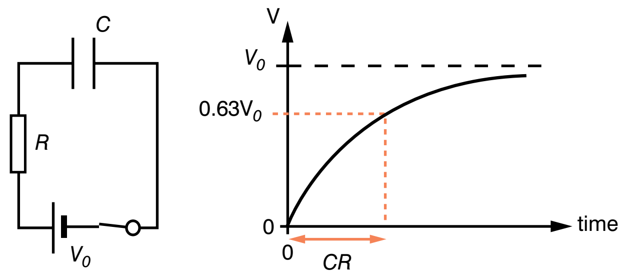


***Figure 2.44** Capacitor charging through a resistor.*

This time when the time constant has elapsed, the capacitor is almost two-thirds full. Such a circuit can be used as the basis for a timing circuit. When the voltage across the capacitor reaches a certain value, it causes something else to occur – a light to come on, for instance. If a control is provided by which you can adjust the time delay, you will be altering a variable resistor to change the value of *R*.

Dielectric materials

A **dielectric** is the electrically insulating material between the metallic plates of a capacitor that increases the capacitance of the capacitor (so a greater charge can be stored at a given voltage). The advantage of using a dielectric is that it stops the two charged plates coming into contact with each other.

The dielectric in a capacitor is often a solid material with high **permittivity** (permittivity is a measure of how an electric field affects, and is affected by, a dielectric medium, and relates to a material's ability to transmit (or 'permit') an electric field).

KEY WORDS

dielectric *electrically insulating material*

permittivity *a measure of how an electric field affects, and is affected by, a dielectric medium*

Challenge...

Work in a small group to derive the given formulae for capacitors in series and in parallel.

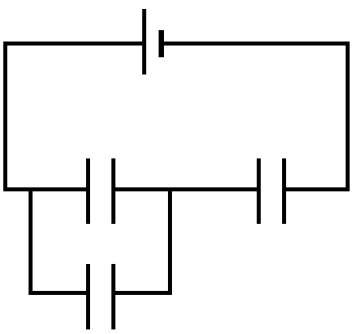


Figure 2.46 Combining parallel and series capacitors.

Activity 2.20: Charging and discharging series and parallel combinations of capacitors

Repeat Activity 2.19, but this time use combinations of capacitors in series and in parallel. Add a resistor to the circuit. Investigate the variation in discharge times. Try to explain your observations.

Combination of capacitors in series, parallel and series–parallel

Sometimes you may wish to combine two or more capacitors, to make one of a different value. The formulae are as follows.

Capacitors in parallel (as shown in Figure 2.45a)

$C_{total} = C_1 + C_2$

Capacitors in series (as shown in Figure 2.45b)

$$\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2}$$

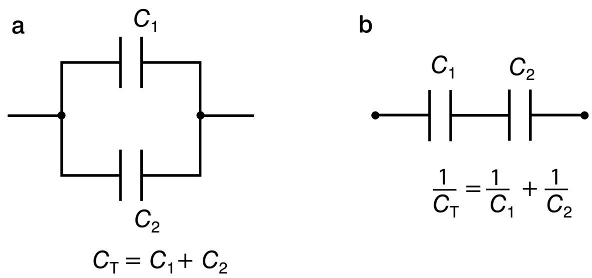


Figure 2.45 Parallel capacitors and series capacitors.

It is also possible to organise capacitors as shown in Figure 2.46, in which case you work out the capacitance of the parallel combination first and then take that capacitance as C_1 in the formula for series capacitors.

Worked example 2.10

Two capacitors of values 4 μF and 2 μF are placed a) in parallel, b) in series. Draw a diagram to show these circuits. Work out the effective capacitance of each combination.

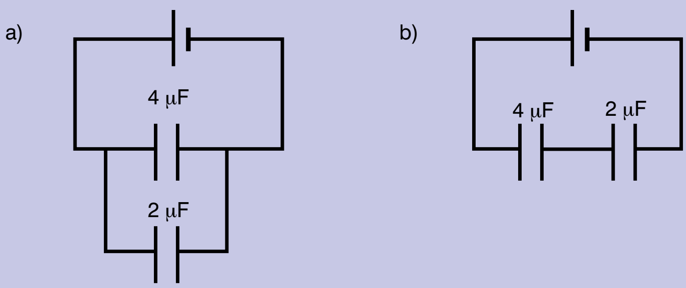


Figure 2.47 Capacitors in parallel use and series use.

a) Capacitors in parallel use $C_{total} = C_1 + C_2$

$C_{total} = 4\ \mu F + 2\ \mu F = 6\ \mu F$

b) Capacitors in series use

$$\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2}$$

$$\frac{1}{C_{total}} = \frac{1}{4\ \mu F} + \frac{1}{2\ \mu F} = \frac{3}{4}\ \mu F$$

So $C_{total} = \frac{4}{3}\ \mu F$

Parallel plate capacitor and the factors that affect its capacitance

A parallel plate capacitor has two plates that are parallel to each other. If a capacitor has two plates each of area A separated by a distance d, then it is possible to calculate what its capacitance C will be by using the relationship:

$C_0 = \frac{\varepsilon_0 A}{d}$ if the distance d is not filled by any dielectric material

If there is a dielectric material,

$$C = \frac{\varepsilon A}{d}$$

The symbol ε is a constant that varies according to the dielectric which is used to separate the plates.

Looking at the expression, in addition to a suitable dielectric, what are needed for a large capacitance are two plates of large area placed exceedingly close together. One way to achieve this is to use a very long strip of thin metal foil for each of the plates. The two strips are separated by a similar-shaped strip, perhaps of thin waxed paper (which acts as the dielectric). The assembly is then rolled up tightly to fit into a plastic case with two leads, one going to each plate.

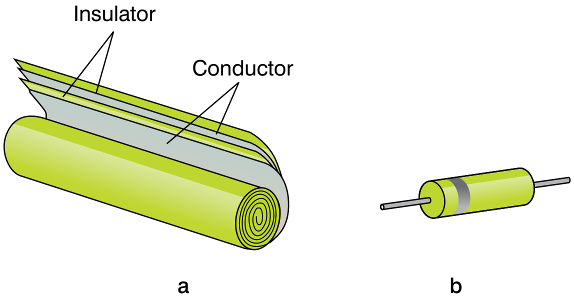


Figure 2.48 *Parallel plate capacitor.*

Another type of capacitor is the electrolytic capacitor, where the three layers are deposited by an electroplating technique, which helps to make them very thin – so the plates are close together and the whole device is compact. (If you wonder how you can deposit an insulator by electroplating, the answer is you cannot. The dielectric has to be slightly conducting, which means that the charge will soon leak off the capacitor. With some applications in electronics the charge needs be stored for only a brief interval, in which case that is not important.)

Activity 2.21

If possible, look for the capacitors in a broken radio. See if you can find the dielectric material.

The electric potential energy stored in a capacitor

If a capacitor holds a charge Q at a potential V, then the energy it contains is given by:

Energy $= \frac{1}{2}QV$

An alternative form of the same expression, including $C = \frac{Q}{V}$ as well, is:

Energy $= \frac{1}{2}CV^2$

Worked example 2.11

A 20 μF capacitor is charged to 12 V. How much energy does it contain?

Energy = $\frac{1}{2}CV^2$ =

$\frac{1}{2} \times 20 \times 10^{-6} \times 122$ =
1.4×10^{-3} J (1.4 mJ)

This is only a tiny amount of energy. Clearly capacitors are not suitable for storing large amounts of energy.

However, a capacitor can be useful for the flash circuit in a camera. The flash needs a big current just for an instant – not much charge is needed in total, but it must all be supplied quickly. The little battery in the camera cannot supply such a large current, but what it does is to slowly fill up a capacitor; it is the capacitor then that empties to generate the flash for an instant. The capacitor is behaving rather like the tank in a flushing toilet. A small water pipe is used to slowly fill up the cistern, then when it flushes the whole tankful empties in a very short time to give a momentary rapid flow of water.

Why is the energy not given by QV? After all, we have Q coulombs of charge being stored at a potential of V volts for each coulomb – so why the extra factor of a half?

To see the answer, it is necessary to consider what happened when the charge first went on to the capacitor, a little at a time. The first piece of charge went on to an uncharged capacitor, and so did not have to be 'lifted' through any volts. Successive charges had to be 'lifted' through an increasing voltage, but it was only the final piece that had to be 'lifted' through the full V volts.

Applying those ideas:

the energy stored on the capacitor =

the total charge × the *average* voltage through which it had to be 'lifted' (which is just $\frac{1}{2}V$)

$= Q \times \frac{1}{2}V = \frac{1}{2}QV$.

It is a bit like building a brick wall which is to be 2 m high. The first bricks have to be lifted through no height at all, and it is only the final ones that have to be lifted through the full 2 m. The average brick has to be lifted by 1 m.

Capacitors in everyday life

In direct current circuits (see Unit 3) a capacitor provides a break in the circuit. Apart from a momentary charging current when you switch on, nothing will happen.

In alternating current circuits (see Unit 3) capacitors do useful things, however.

Look at Figure 2.49, which shows a capacitor in a circuit that is being powered by a 50 Hz alternating voltage supply. This means than the voltage rises to a maximum, dies away, builds up to a maximum in the other direction, then dies away again – and repeats this cycle 50 times every second.

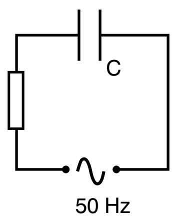


Figure 2.49 *Capacitor powered by 50 Hz alternating voltage supply.*

The capacitor is continually charging up, emptying, charging up the other way round then emptying again – and does this 50 times every second. As far as the power supply and the rest of the circuit are concerned, it is just as if there was a complete circuit.

To sum up, therefore, capacitors block direct currents but appear to allow alternating currents to pass. That is one important use in electronic circuits.

In a public address system, for example, the signal is an alternating current. This has to be passed on from one amplifier stage to the next (see Figure 2.50).

The output of the first stage is at a different voltage to the input of the next stage. Just joining them with a wire would have caused both stages to stop working. The capacitor provides the solution.

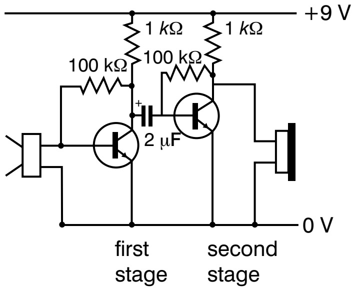


Figure 2.50 *Alternating current in a public address system.*

Summary

- A capacitor is a small device designed to store more charge at a lower potential. The commonest way of doing this is to use two parallel plates, a tiny distance apart and separated by an insulator.
- If we place some charge on an insulated metal sphere, the sphere's voltage will rise; put some more charge on, and its voltage will rise further.
- Presumably a larger sphere will hold more charge before its potential has risen by 1 V. We measure this by using a quantity called capacitance, C. If a charge Q results in a rise of V in the potential, we define the capacitance of the sphere by:
  $C = \frac{Q}{V}$
- This is the charge needed for each volt rise in the sphere's potential. The units will be coulombs per volt, C/V, which we call farads, F.
- You can apply the definition of capacitance to solve numerical problems.
- The circuit symbol to represent a capacitor is as shown in Figure 2.51.
- You can use this symbol to draw an electric circuit diagram for a simple capacitor and series and parallel connections of two or more capacitors.
- When a capacitor discharges through a resistor (or a bulb), the capacitor acts like a battery to drive a current through the resistor, but unlike a battery its voltage drops rapidly as its charge drains away.
- When a capacitor is charged through a resistor we have the opposite to the discharge situation. Here the capacitor charges rapidly at the start, but then continues at an ever-declining rate. It is a kind of 'upside down' exponential curve. When the voltage between the plates of the capacitor equals that of the battery, the charging ceases.
- A dielectric is the electrically insulating material between the metallic plates of a capacitor and increases the capacitance of the capacitor (so a greater charge can be stored at a given voltage). The advantage of using a dielectric is that it stops the two charged plates coming into contact with each other. The dielectric in a capacitor is often a solid material with high permittivity (permittivity is a measure of how an electric field affects, and is affected by, a dielectric medium, and relates to a material's ability to transmit (or 'permit') an electric field).
- Sometimes you may wish to combine two or more capacitors, to make one of a different value. The formulae are as follows:

  Capacitors in parallel (as shown in Figure 2.51a)

  $C_{total} = C_1 + C_2$

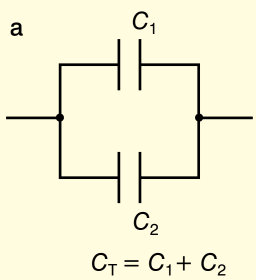


Figure 2.51a

- Capacitors in series (as shown in Figure 2.51b)

  $\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2}$

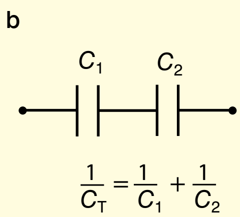


Figure 2.51b

Summary

- It is also possible to organise capacitors as shown below, in which case you work out the capacitance of the parallel combination first and then take that capacitance as C_1 in the formula for series capacitors.

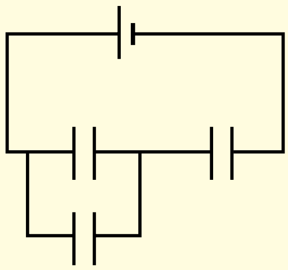


Figure 2.52

- You can use the above expressions to solve problems on combination of capacitors.
- A parallel plate capacitor has two plates that are parallel to each other. If a capacitor has two plates each of area A separated by a distance d, then it is possible to calculate what its capacitance C will be by using the relationship:

$$C = \frac{\varepsilon A}{d}$$

  The symbol ε is a constant that varies according to the dielectric which is used to separate the plates.

- From the expression, the factors that affect the capacitance of a parallel plate capacitor are the dielectric material, the area of the plates, and the distance between the plates.
- You can calculate the capacitance of a parallel plate capacitor using the above formula.
- The electric potential energy stored in a capacitor can be found using

  $\text{Energy} = \frac{1}{2}QV$

  or $\text{energy} = \frac{1}{2}CV^2$
- You can calculate the energy stored in a capacitor using one of these formulae.
- Some uses of capacitors in everyday life include in a camera flash, to provide a break in a direct current circuit, to allow alternating currents to pass, to link two stages of an amplifier, and as the basis for a timing circuit.

Review questions

1. Describe the structure of a simple capacitor.
2. Define the term capacitance and its SI unit.
3. A 2.0 μF capacitor is charged from a 6.0 V battery.
   a) How much charge will it hold?
   b) How much energy is stored on it?
   c) Where has that energy come from?
   d) If the capacitor is then discharged, where does the energy go to?
4. a) Take a 2.0 μF and a 5.0 μF capacitor and connect them in parallel. Draw the circuit diagram. What is the capacitance of the combination?
   b) Repeat part a), but this time connect capacitors in series.
5. Explain the charging and discharging of a capacitor.

6. Define the term dielectric and explain what is meant by a dielectric material.
7. Explain the effect of inserting a dielectric in the gap between the plates of a parallel plate capacitor.
8. Define a parallel plate capacitor and describe the factors that affect its capacitance.
9. You want to make a 5.0 μF capacitor. For the dielectric you use waxed paper of thickness 0.4 mm and $\varepsilon = 7.1 \times 10^{-12}$ F/m. What area plates will be required?
10. State an expression for the electric potential energy stored in a capacitor.
11. A photographic flash unit has a 100 μF capacitor which is charged up by a small 22.5 V battery. How much energy may be discharged into the flash bulb?

End of unit questions

1. Describe how you can test how charged bodies attract or repel one another.
2 . Why is the inside of a car a safe place to shelter in a storm?
3. a) Find the force acting on a charge of 4 C in an electric field of 5 N/C.
   b) State Coulomb's law.
   c) State the similarities between Coulomb's law and Newton's law of universal gravitation.
4. What is an equipotential?
5. a) What is a capacitor?
   b) State some uses of capacitors in everyday life.
   c) Two capacitors of 5 μF and 10 μF could be arranged either in series of in parallel. Which arrangement gives the highest capacitance? Explain your answer.
   d) If two capacitors, A and B, have the same area and dielectric constant, but the distance between the plates of A is twice the distance between the plates of B, which has the higher capacitance? Explain your answer.
6. An electron travels at a horizontal velocity of 5×10^6 m/s in a vacuum between two charged plates 100 mm long and 20 mm apart, to which a p.d. of 30 V is applied. The charge on an electron is 1.5×10^{-19} C and the mass of an electron is 9.1×10^{-31} kg.
   a) What size force will the electron experience while it passes through the plates?
   b) What acceleration will this give to the electron?
   c) If the positive plate is the lower one, in which direction will the electron accelerate?
   d) How long will it take for the electron to pass through the plates?

Current electricity

Unit 3

Contents

Contents

3.1 Electric current

By the end of this section you should be able to:

- Define electric current and its SI unit.
- Explain the flow of electric charges in a metallic conductor.
- Calculate the number of electrons that pass a point in a given length of time when the current in the wire is known.

KEY WORDS

ampere *SI unit of electric current*

Electric current and its SI unit

An electric current is a flow of charge. Comparing it with water, a small current is like a trickle passing through a pipe; a really large current is like a river in flood.

The rate of flow of electric charge – that is, the electric current – is measured in amperes (A). The **ampere** is one of the fundamental units of the SI system. This means that the size of the ampere is not fixed in terms of other units: we simply compare currents to a 'standard ampere'.

Household appliances, such as toasters and kettles, run on a current of a few amperes. An ampere is quite a sizeable flow of charge; however, especially in electronic circuits, we also often deal in milliamperes (mA, thousandths of an ampere, 10^{-3}A) or even microamperes (μA, millionths of an ampere, 10^{-6} A).

If charge flows at a rate of one ampere, and continues to flow like that for a second, then the total amount of charge that has passed is one coulomb. This is how the size of the coulomb is fixed: in terms of the ampere and the second.

In practice it is probably best to picture coulombs of charge flowing at a rate of so many amperes. A current of 3 A, for example, is a flow rate of 3 coulombs of charge every second (3 C/s). With that

Worked example 3.1

120 C of charge passes in 1 minute. What is the current?

$$I = \frac{Q}{t} = \frac{120\ \text{C}}{60\ \text{s}} = 2\ \text{C/s} = 2\ \text{A}$$

KEY WORDS

galvanoscope *an instrument for detecting the presence of an electric current*

electron *a negatively charged particle that orbits round the nucleus of an atom*

ion *an atom bearing unequal numbers of electrons and protons*

conduction electrons *electrons in the conduction band of a solid, free to move under the influence of an electric field*

insulator *a material that resists the flow of electric charge*

current, therefore, it should be obvious that in 10 seconds a total of 30 C of charge will pass, or that to supply 12 C of charge the current must flow for 4 seconds.

You are strongly advised to understand the previous paragraph, rather than just remember. When you need the formula linking amperes, coulombs and seconds, however, it is:

$\boldsymbol{Q = It}$ Units

Q coulombs (C)
I amperes (A)
t seconds (S)

In the formula, Q stands for the quantity of charge that passes when a current I flows for a time t.

Worked example 3.2

How long will a current of 5 A take to pass 100 C of charge?

$$Q = It, \text{ so } t = \frac{Q}{I} = \frac{100\ \text{C}}{5\ \text{A}} = 20\ \text{s}$$

Activity 3.1: Making a current tester

You can use a compass needle to make a simple **galvanoscope** (an instrument that detects the presence of an electric current). Simply wrap a few turns of thin insulated copper wire around a compass, connect the ends as shown in Figure 3.1 to a 1.5 V cell and a 1.5 V light bulb. When the wires are connected as shown, the bulb will light and the compass needle will move around in one direction. If you want to investigate this further, disconnect the cell, turn it round and reconnect it the other way round. The needle should move in the opposite direction.

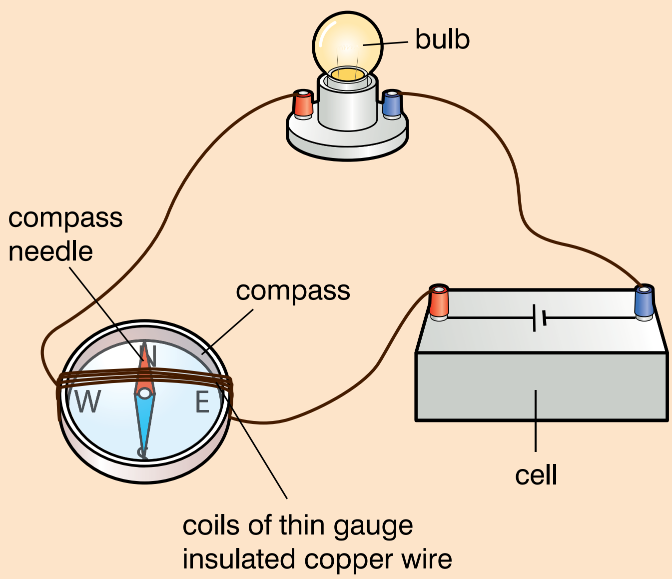


Figure 3.1 *Current tester.*

What are the charges that flow round a circuit?

To answer this question we must first look in rather more detail at the structure of an atom. All its positive charges are located in the central part, the nucleus. Each chemical element has a different number of positive charges in its nucleus, from one to nearly 100. Take copper as an example: it has 29 positive charges in the nucleus.

This means that an uncharged copper atom must have 29 negative electrons as well. They orbit round the nucleus, and each one has its own path. The first two electrons orbit in the innermost shell, the next eight fill the second shell and the following 18 just complete the third shell. That leaves a solitary electron as the first member of a new fourth shell.

It is this electron that accounts for the behaviour of copper as a conductor. It is comparatively easy to remove this electron, changing an uncharged copper atom into what we call a copper **ion** with an overall single positive charge (Figure 3.2).

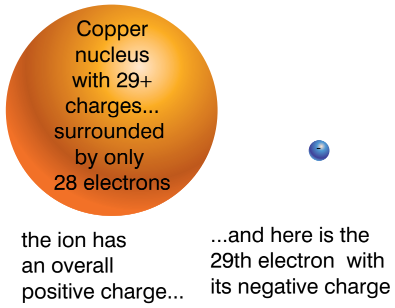


Figure 3.2 *The copper atom can be changed into a positive copper ion plus a negative electron.*

In a copper wire the atoms are packed close together just as in any other solid. More precisely, what are packed together are the positive ions; the complete atoms except for those single outer electrons. They are there as well, so the metal as a whole is uncharged.

In each copper atom 28 of the electrons are still firmly bound in orbit around their nucleus, fixed in its place in the solid. The 29th electrons we call the **conduction electrons**. They remain trapped within the metal as a whole, but otherwise are free to drift about inside it.

They are the charges that move when an electric current flows down the wire. Being negative they will be repelled from the cell's negative terminal and attracted to the positive one (Figure 3.3).

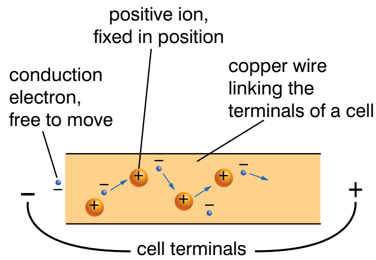


Figure 3.3 *The unattached electrons in the copper wire move when it is connected to the terminals of a cell.*

All metals have these extra one or two outer electrons, so they will all conduct electricity. With most other elements, every electron without exception is tightly bound to its own nucleus, so most other elements are **insulators**.

Conventional current

Electric currents have been investigated since the first cell was invented around 1800. It soon became obvious that charges were flowing round the circuit, but there was no way of telling whether they were positive charges going one way or negative charges going the other.

It was agreed to picture the flow as positive charges repelled from the positive plate, and this was referred to as conventional current.

When the electron was discovered in 1895, it was realised that the guess had been wrong. However, so firmly fixed was the idea that even today we still mark on our circuits the 'conventional current' flowing from the positive pole of the cell to the negative.

You must understand, however, that the electron flow is really negative charge going the opposite way.

Activity 3.2: Testing conductivity

Not all materials have conduction electrons in their structure, so not all materials will conduct electricity. Use your simple galvanoscope (from Activity 3.1) to test which materials conduct electricity and which do not.

Connect your galvanoscope to a copper rod as shown in Figure 3.4a. Observe that the compass needle moves, showing that the rod conducts electricity, allowing an electric current to flow. Connect your galvanoscope to a glass or plastic rod, as in Figure 3.4b. Observe that the compass needle does not move, showing that the rod does not allow an electric current to flow. Repeat this for a selection of different materials of the same size, and rank your materials in order of the size of the deflection of the compass needle.

You will see that some materials conduct electricity better than others, while some materials do not conduct electricity at all.

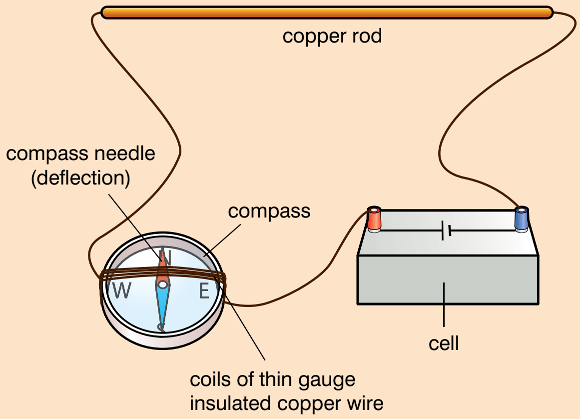


Figure 3.4a

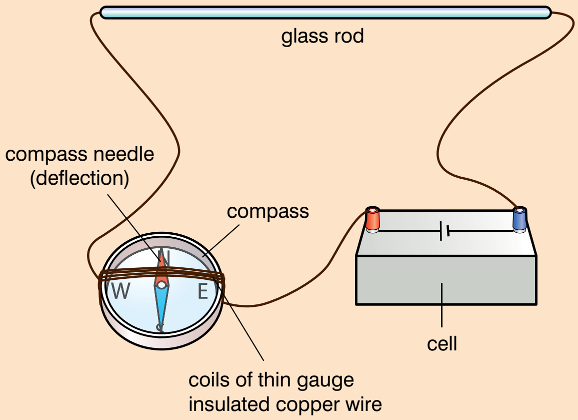


Figure 3.4b

KEY WORDS

drift speed *the average speed that an electron attains in an electric field*

Calculating the number of electrons that pass a point at a given length of time when the current in the wire is known

The wire shown in Figure 3.5 carries a current of I amperes. There are n free electrons per cubic metre of wire. Each electron carries a charge of e coulombs, the cross-sectional area of the wire is A square metres and the average **drift speed** of the electrons in this material is v metres per second.

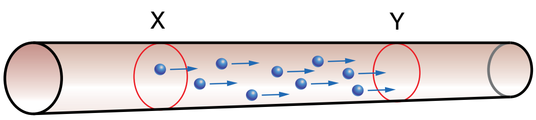


Figure 3.5

Assuming that it takes t seconds for an electron to pass from X to Y (and using the formula distance = speed × time), distance X–Y is vt metres. The volume of wire between X and Y is therefore Avt cubic metres.

The number of electrons between X and Y is therefore $nAvt$.

As each electron carries a charge of *e* coulombs, the total charge passing point Y in *t* seconds is *nAevt*.

The current, *I*, is the charge per second

$$I = \frac{nAevt}{t}$$

or $I = nAev$

As the charge carried by an electron is known to be 1.6×10^{-19} C, the number of electrons passing through a wire can be calculated if the current in the wire is known.

Worked example 3.3

A current of 5 A is flowing through a 2 mm diameter wire. The drift speed of electrons in the wire is 10^{-5} m/s and the charge on the electron is 1.6×10^{-19} C. Calculate the number of electrons passing a point in the wire in a second.

The cross sectional area (*A*) of the wire is πr^2.

$$A = \pi \times \frac{d}{2} \times \frac{d}{2} = \pi \times 0.001 \times 0.001 = \pi \times 10^{-6}\ \text{m}^2.$$

Using $n = \dfrac{I}{Aev}$

$$\text{number of electrons} = \frac{5}{\pi \times 10^{-6} \times 1.6 \times 10^{-19} \times 10^{-5}}$$

$= 9.947 \times 10^{29}$

Study of electric charges in a metallic conductor

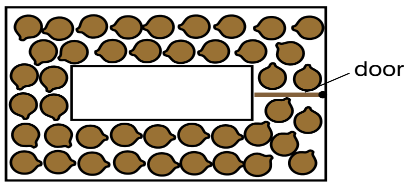


Figure 3.6

In Activities 3.1 and 3.2 you detected an electric current. Where does this electric current keep coming from and where does it keep going to? To start investigating the answer to this, look at Figure 3.6 which shows the top view of a corridor that ends up where it started. The whole place is filled with people who form a kind of endless queue. The door is shut, which means that nobody in the queue can move forward. Once the door is opened (once the circuit is switched on), all the people in that corridor can immediately start moving. They can keep circulating round and round the corridor, thus forming a continual current.

Of course when the door was opened the people filling the corridor found themselves free to move round the circuit, but they did not have to do so. Likewise charges will not circulate to give a current unless there is something that keeps 'pumping' them round. That is the job of the cell, as we shall see below.

KEY WORDS

electrochemical cell *a device capable of deriving electrical energy from chemical reactions*

electrodes *conductors used to make contact with part of a circuit*

electrolyte *a solution that conducts electricity*

electromotive force *a source of energy causing current to flow in an electrical circuit*

e.m.f. *electromotive force*

dry cell *electrochemical cell containing electrolyte in the form of a paste*

polarisation *the formation of a film of hydrogen gas on the positive plate of a dry cell*

Electric energy from chemicals

In Activities 3.1 and 3.2 a flow of electric current is detected by a galvanoscope. What causes this current to flow around the circuit?

An electrochemical cell is one device that makes current flow round a circuit. Such cells have two **electrodes** made of two different **conductors**, an **electrolyte** solution (a solution that conducts electricity) which reacts with the electrodes, and a conductive wire through which electrons can flow.

Activity 3.3: Make your own electrochemical cell

Roll a lemon, orange, grapefruit, or other citrus on a firm surface to break the internal membranes.

Insert two rods – one copper, one zinc – into the fruit and connect as shown in Figure 3.7. Observe the size of the deflection of the needle on the instrument.

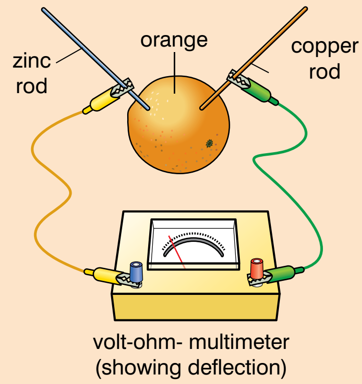


Figure 3.7

The electrodes in this cell are the copper rod and the zinc rod. The electrolyte here is the juice inside the orange.

Repeat the experiment using rods made from different materials and observe the deflection of the needle on the instrument. Do some pairs of materials produce a greater deflection than others?

The volt

A volt is defined as the potential difference (p.d.) required to produce a current of 1 A in a circuit with a resistance of 1 ohm (or Ω). It is named after the Italian scientist Alessandro Volta (1745–1827).

Some cells are more powerful than others. An ordinary torch cell (usually referred to as a battery) is rated at 1.5 volts. We call this the **electromotive force** of the cell, usually shortened to **e.m.f**. It is the cell that pumps the charge round the circuit, and for the time being it is sufficient to think of the e.m.f. as its 'electrical pumping strength'. All other things being equal, the greater the e.m.f. of the cell the greater the current that it will drive round a circuit. You will have got the right idea when you do not like to hear people talking about volts flowing round a circuit. Volts do not flow. They cause coulombs of charge to flow round a circuit at a rate of amperes.

Sources of electricity

There are many different types of cell. Figure 3.8 illustrates the principles of one common form, which we call a **dry cell**. Experiment shows that the carbon plate is the positive one, the zinc is the negative, and the e.m.f. is very close to 1.5 V.

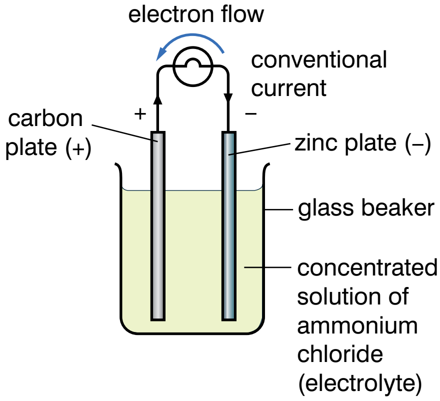


Figure 3.8 *The principle of a common sort of cell.*

Such a cell is about as portable as a glass of water, and we now see how it has been converted into the tough flashlight cell (commonly known as a battery) we are familiar with.

The negative zinc plate is shaped to form the casing of the cell (Figure 3.9). The carbon runs down the centre in the form of a rod (with a brass cap on the top, because carbon is easily damaged). The concentrated solution of ammonium chloride that has to go into the space between them is made much less runny by forming it into a paste or jelly. It is for this reason that the design is called a dry cell.

This leaves the complete cell as shown in Figure 3.10.

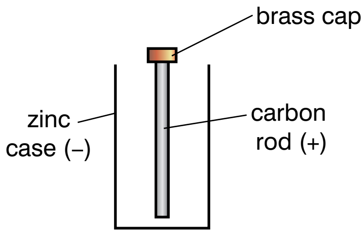


Figure 3.9 *The zinc casing forms the negative plate.*

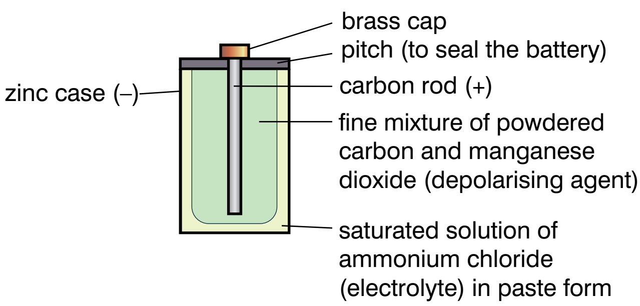


Figure 3.10 *The complete dry cell.*

The positive plate is both the carbon rod down the centre and the thick layer of powdered carbon that surrounds it.

Mixed in with the powdered carbon is manganese dioxide, another chemical in powder form. While the cell is giving a current it suffers from **polarisation**: the formation of a film of hydrogen gas on the positive plate which 'clogs' the cell up. The manganese dioxide is there to deal with this problem by supplying oxygen to convert this unwanted hydrogen to water. Sometimes the mixture of black powders is held in place round the central carbon rod with a cloth bag, but in other batteries it is just the 'stiffness' of the ammonium chloride paste that keeps it there.

KEY WORDS

primary cells *electrochemical cells that transform chemical energy directly to electrical energy*

secondary cells *electrochemical cells that have to be charged up by passing an electric current through them*

As the cell runs down, the zinc casing gets thinner and thinner, and when it gets very old the ammonium chloride may seep out from holes in the casing as a white corrosive paste.

There is nothing unique about carbon, zinc and ammonium chloride solution: any combination of two different metals (or one metal and carbon) placed in a solution of an electrolyte will produce a voltage. Other combinations are available, such as the alkaline cell, which uses potassium hydroxide for the electrolyte.

Primary cells and secondary cells

Cells that produce their voltage directly from chemical energy stored in the ingredients which make them up are known as **primary cells**. When all their chemicals have reacted, you have to buy a new one. The dry cell is one of these.

There are also **secondary cells**, which have first to be charged up by forcing a current 'backwards' through them. The commonest types of these are the increasingly popular rechargeable cells and the lead–acid accumulators that make up the battery in a car.

Activity 3.4: The direction of current flow

Part A

Connect the cell, motor and switch (open) as shown in Figure 3.11 on the edge of a table or bench. Clamp the motor to the edge of the table. Attach weights to the spindle of the motor, to hang over the edge of the table or bench. Close the switch. The motor will turn and the weights will move. Observe the movement of the weights; one weight will rise and one weight will fall.

Part B

Attach the cell the opposite way round. Close the switch and observe the movement of the motor and the weights. (x rises, y falls)

You will see that the motor turns in the opposite direction in Part B, causing the weight which rose in Part A to fall in Part B, and the one which fell in Part A to rise in Part B.

This demonstrates that the electrons in the circuit flowed in one direction in Part A and in the opposite direction in Part B.

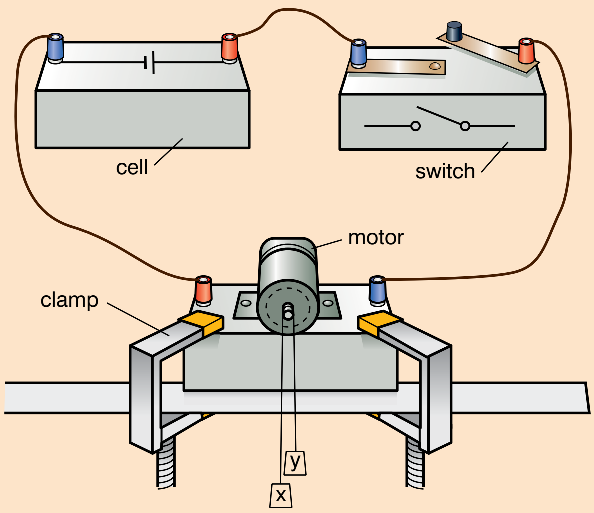


Figure 3.11

The 'standard ampere' – a note to turn back to when you have covered more physics

You may be puzzled as to how anyone can keep a standard ampere somewhere. That is not possible, of course, but we do have a way of defining it. (There is no need for you to remember the details.)

Two parallel wires each carry the same current (see Figure 3.12) and each current produces a magnetic field.

The magnetic field produced by one current-carrying wire acts on the other current-carrying wire to cause the **motor effect**, which pushes the wires apart.

The bigger the current in the wires, the greater the force pushing them apart.

If the wires are 1 metre apart and the push reaches 2×10^{-7} newtons on each metre length, then we say that each of the currents is one ampere.

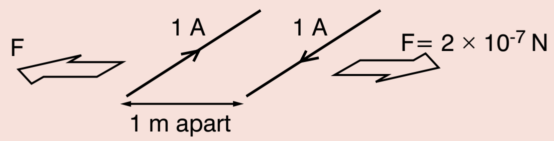


Figure 3.12

KEY WORDS

motor effect *when a current-carrying wire passes at right angles through a magnetic field, a force is exerted on the wire*

thermoelectricity *electricity produced when two metals of different temperatures are joined in a circuit*

Electric energy from heat

We have seen how electric energy can come from chemicals and we now see that it can also come from heat; this is called **thermoelectricity**.

If two different metals are joined in a circuit, and if one junction between the metals is hotter than the other, a very small voltage (a few millivolts, mV) is generated. This effect is known as the Seebeck effect after the German physicist Thomas Seebeck (1770–1831) who discovered it.

The circuit must be made from two different metals, any two, although some pairs may be better than others.

The thermocouple thermometer

The Seebeck effect is used in the thermocouple thermometer to measure temperature. It consists of a circuit made from any two different metals and includes a sensitive meter to measure small currents (Figure 3.13). If one of the junctions between the metals is hotter than the other, a small voltage is generated, which drives a tiny current round the circuit. The larger the temperature difference, the greater the current.

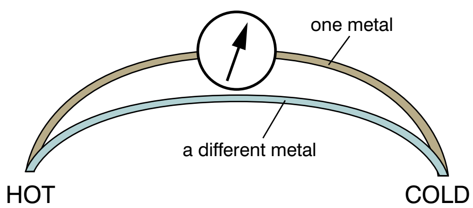


Figure 3.13 *The thermocouple thermometer.*

When using a thermocouple thermometer, it is usual to hold one junction at 0 °C by keeping it immersed in melting ice. The other junction acts as the probe to investigate the temperature to be measured, as in Figure 3.14.

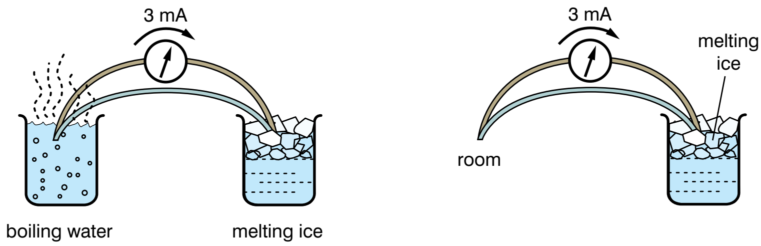


Figure 3.14 The thermocouple thermometer in use.

In the first part of Figure 3.14, a temperature difference of 100 °C between boiling water and melting ice shows a current of 10 mA (milliamperes, thousandths of an ampere).

Each 1 mA therefore indicates a difference of 10 °C.

Thus when the probe is at room temperature and a current of 3 mA is measured, the junction at room temperature must be 30 °C hotter than the cold junction, which is at 0 °C. Room temperature is therefore 30 °C.

Worked example 3.4

One junction of a thermocouple thermometer is immersed in melting ice and the other in boiling water. A current of 20 mA is recorded. The thermocouple thermometer is then used to measure the temperature of a liquid. One junction is immersed in melting ice and the other in the liquid whose temperature is to be measured. A current of 5.5 mA is recorded. What is the temperature of the liquid?

In the first, calibration, measurement, 20 mA represents a temperature difference of 100 °C. Therefore 1 mA represents a temperature difference of 5 °C. In the temperature measurement, the temperature difference between the junctions is 5.5 × 5 = 27.5 °C.

The junction immersed in melting ice is at 0 °C. The temperature of the liquid being measured is therefore

0 + 27.5 = 27.5 °C.

Activity 3.5: Making a thermocouple thermometer

Create thermocouple junctions at both ends of a section of iron wire by twisting the ends together with copper wires.

Place one copper–iron junction in a beaker with ice water and leave the other junction outside. The two remaining ends of the copper wires should be connected to a sensitive galvanometer.

Heat the exposed junction with a Bunsen burner or match and record the current.

Does the current increase or decrease if the heat source is removed? Is the change in current immediate? Discuss these questions.

Summary

- Current is a flow of electric charge.
- Coulombs of charge flow at a rate of amperes.
- We need a voltage supply to cause the charge to circulate round the circuit.
- A primary cell uses the chemicals in it to supply electrical energy; a secondary cell has to be charged up first.

Review questions

1. Charge flows along a wire at rate of 3 A. How many coulombs of charge will pass a given point in the circuit in:
   a) 1 s
   b) 12 s
   c) 2 minutes?
2. A very sensitive ammeter records a current of 30 μA. How long will it take before 6 μC of charge has flowed through it?
3. In an experiment to copperplate a coin, a current of 100 mA is passed through a solution of copper sulphate. If 30 C of charge must pass before the coin has enough copper deposited on it, for how long must you leave the current switched on?
4. a) A single electron has a charge of 1.6×10^{-19} C. How many electrons must pass round the circuit of question 5 to achieve the 30 C?
   b) For every two electrons that arrive at the coin which is being plated, one copper atom is deposited. If the mass of a copper atom is 1.1×10^{-25} kg, what is the total mass of the copper which now plates the coin?
5. What name do we give to the unit 'ampere second' (A s)? What physical quantity would be measured in such a unit?
6. You have two batteries, a large one and a tiny one, each consisting of a single dry cell. In terms of their performance, what would you expect to be the same for the two batteries and what would you expect to be different about them?
7. What is the difference between a primary cell and a secondary cell?
8. What temperature does the thermocouple in Figure 3.15 indicate for:
   a) air
   b) liquid B
   c) liquid C?

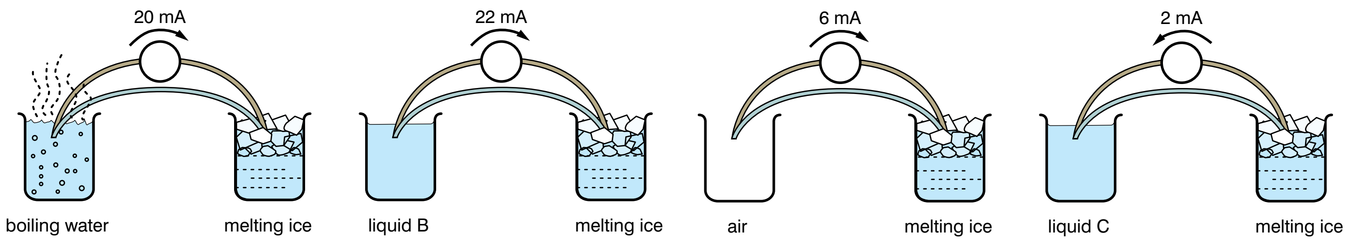


Figure 3.15

9. A thermocouple junction will respond to changes in the temperature of its surroundings more quickly than a mercury-in-glass thermometer would. There are two reasons for this:
   a) the junction is smaller than the bulb full of mercury
   b) the junction is not encased in glass.

   Explain why each of these factors helps the thermocouple to respond quickly.
10. A metal wire is uncharged. Explain how it is possible for a current to flow through it.

3.2 Ohm's law and electrical resistance

By the end of this section you should be able to:

- Describe factors affecting the resistance of a conductor.
- Write the relationship between resistance R, resistivity ρ, length l and cross-sectional area A of a conductor.
- Calculate the resistance of a conductor using the formula $R = \rho l/A$.
- Find the relationship between resistivity and conductivity.
- Construct and draw an electric circuit consisting of a source, connecting wires, resistors, a switch and a bulb using their symbols.
- Explain why an ammeter should be connected in series with a resistor in a circuit.
- Explain why a voltmeter should be connected in parallel across a resistor in a circuit.
- Do experiments using an ammeter and a voltmeter to investigate the relationship between current and p.d. for metallic conductors at constant temperature.

Electrical resistance

The size of the current a cell will pump round a circuit depends on two things. One is the electromotive force, or e.m.f., of that cell, measured in volts. As we shall see in Section 3.4, adding a second cell **in series** (the current passes through one cell then the other) with the first makes a battery of cells, which has twice the e.m.f. of a single cell, and the charge will be pumped round the circuit at twice the rate. In other words, the current will double.

The second factor determining the current is the circuit through which the cell must drive the charge. Is it an easy circuit made of thick pieces of a good conductor, or is it a more difficult circuit consisting of thin wire made from a metal that does not conduct electricity so well? It is this second factor we are going to consider now. Does the circuit have a low **resistance**, or does it have a high resistance?

We specify the resistance of a circuit by the number of volts of 'battery power' we would need to get a current of 1 A to flow round it. A low resistance circuit might need only 2 V per ampere (2 V/A), say. This suggests that a 1 V battery would produce a flow rate of 0.5 A round it, or that to get a current of 3 A going you would need to provide an e.m.f. of 6 V. A less easy circuit might have a resistance of 200 V/A; in other words, as many as 200 V would be needed to establish a current of just 1 A. It is important to study the behaviour of resistors in electrical circuits.

KEY WORDS

in series *wiring an electrical circuit so that there is only one path for the current to take between any two points*

resistance *the opposition to a flow of current in an electrical circuit*

Measuring the resistance of a resistor

The principle is simple: apply a voltage across the resistor, and measure the size of the resulting current. A suitable circuit is shown in Figure 3.16 where *R* is the resistor to be measured.

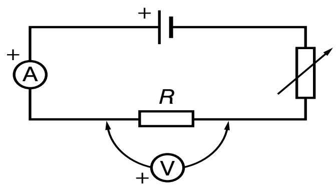


Figure 3.16 *Measuring the resistance of a resistor.*

Figure 3.16 also shows the symbol of a variable resistor (sometimes called a 'rheostat'). By moving a slider or rotating a knob you can alter its resistance. This in turn will alter the total resistance of the whole circuit, and will therefore control the current drawn from the battery which then passes through the ammeter.

In a lighting circuit a variable resistor would act as a dimmer switch to make the bulb fainter or brighter.

The ammeter may be placed anywhere in the circuit since the same current flows all the way round. The voltmeter is not part of the circuit itself; it is placed alongside to measure the drop in voltage between the two ends of *R*. The variable resistor is not essential, but it enables you to alter the voltage drop across *R* and the current through it so as to get check readings.

If you try this activity for yourself, you should choose a resistor to measure that is not significantly warmed by the current you send through it. A length of resistance wire open to the air should do, but avoid using a light bulb. You will find out why in the next section.

Your results should be recorded in a table like that in Figure 3.17.

p.d. *V* across *R* (volts)	current *I* through *R* (amperes)	$R = \frac{V}{I}$ (ohms)
	Average	

Figure 3.17

KEY WORDS

Ohm's law *the current that flows through a conductor is proportional to the potential difference between its ends*

The ohm (Ω)

The unit for resistance is the ohm (Ω).

It is named after Georg Ohm (1787–1854), a German physicist who was one of the first to investigate how currents flowed in circuits.

The abbreviation for ohm should really be a capital 'O', but that could be confused with a zero. Luckily there is a letter in the Greek alphabet called 'omega', which provides the ideal replacement: the abbreviation for ohms, therefore, is 'Ω', which is a capital omega.

Worked example 3.5

A 12 V car battery is connected to a circuit for which total resistance is 6 Ω. What current will flow?

$$I = \frac{V}{R} = \frac{12}{6} = 2\text{ A}$$

As you repeat the experiment for different values of V and I, you should find that the value for R remains constant. This illustrates **Ohm's law**.

Ohm's law

For a metal wire at a constant temperature, the current that flows through it is proportional to the potential difference (the voltage drop) between its ends.

In other words, if the voltage drop across the wire doubles, charge will flow through it at exactly twice the rate (that is, the current doubles too). Put yet another way, $\frac{V}{I}$ stays constant: the wire's resistance does not change. Notice that this will apply only if the temperature of the conductor does not change.

This can be put into a formula. The current I that is produced when a battery of e.m.f. V is connected to a circuit of resistance R is given by:

$$I = \frac{V}{R}$$

Current (A)	I	ampere
Voltage (V)	V	volt
Resistance (Ω)	R	ohm

To make the current larger you could either increase V by adding another cell or you could reduce the resistance of the circuit.

Worked example 3.6

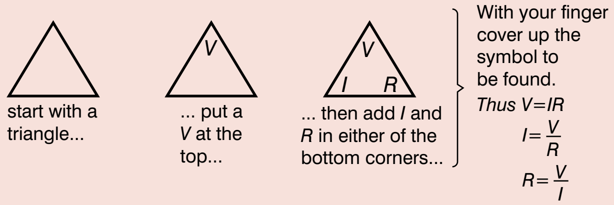
A 12 V battery is connected to a circuit. If the current is 2 mA, what must the resistance of the circuit be? (Notice here that the current is given as 2 mA. Before it can be entered into the formula it must be converted to 0.002 A (or to 2×10^{-3} A.)

$I = \frac{V}{R}$, so rearranging we get $R = \frac{V}{I} = \frac{12}{0.002} = 6000\ \Omega$

A thousand ohms is one kilohm (rather than 'kiloohm'), so the answer could be expressed as 6 kΩ.

A trick to help you with the calculations

If you have had a problem with changing $I = \frac{V}{R}$ into $R = \frac{V}{I}$, there is a method which will help. The diagram shows you how.



Provided you remember that V goes at the top, you will not go wrong.

Worked example 3.7

What voltage battery would be needed to send a current of 3 A round a circuit for which the total resistance is 4 Ω?

Here $I = 3$ A and $R = 4$ Ω.

$V = IR = 3 \times 4 = 12$ V

Factors affecting the resistance of a conductor

Effect of heat on resistance

If you display the current and voltage readings measured when studying a resistor at constant temperature (recorded in a table such as Figure 3.17) in the form of a graph of the current through the resistor plotted against the voltage drop applied across it, you should get a straight line (Figure 3.18). Double the potential difference (p.d.) across the wire, and you will double the current flowing through it.

However, if you use a light bulb as your resistor and take a range of readings such that the filament of the bulb is varied from not even red-hot to brilliantly white-hot, the graph of your readings will form a curve (Figure 3.19). As you increase the voltage and the lamp glows more brightly, the current does not rise as rapidly as expected.

As the filament becomes white hot, its temperature increases by at least several hundred degrees. The resistance of metals rises with temperature, and that is why a hot bulb does not conduct as well as a cool one.

Instead of taking current and voltage readings, it is useful to measure the resistance of a light bulb at as wide a range of currents as possible. At one extreme, readings should be taken when the current is so small that the filament is not even glowing feebly red; at the other extreme the bulb should be brighter than normal.

Use these readings to plot a graph of resistance against current, as in Figure 3.20 on the next page. If the temperature had not changed, the resistance would have been constant (line (a)). As it is, the resistance rises markedly (line (b)).

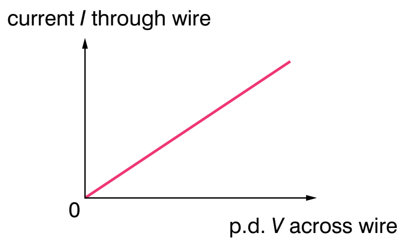


Figure 3.18 *A graph of current against voltage drop for a resistor, R.*

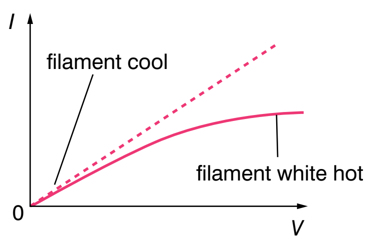


Figure 3.19 *As the temperature of the filament rises, so does its resistance.*

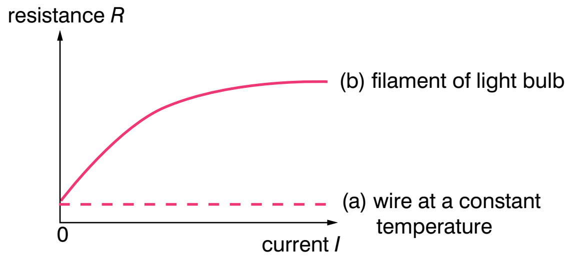


Figure 3.20 Resistance–current graph.

Notice that the line will not pass through the origin. If you continue the curve back so it approaches or equals zero, you will have the resistance of the bulb when it is cold (that is, at room temperature). At its working temperature the resistance of the bulb may well have risen to ten times that value.

The resistance of a metal increases in an approximate straight line if you plot it against its temperature. The increase in the value of the resistance as it warms up depends on:

- The rise in the wire's temperature, ΔT. The greater the rise, the bigger the increase.
- How many ohms of resistance it possesses. Since this will vary with its temperature, we work on the basis of its resistance at 0 °C and call it R_0. The greater the number of ohms at the start, the greater the increase in resistance will be.
- The metal it is made from.

Putting these together, we can say:

the increase in resistance = $\alpha R_0 \Delta T$

Alpha (α) is a constant, which varies from one metal to another. We call it the metal's temperature coefficient of resistance. The units of α have to be K^{-1} (kelvin^{-1}) in order to give an answer that is in ohms. You will probably be working in degrees Celsius rather than kelvin, but there is no need to worry, remember that a rise in temperature of 10 °C is exactly the same as a rise of 10 K.

Worked example 3.8

A copper resistor (α = 4.3×10^{-3} K^{-1}) whose value is 10.0 Ω at 0 °C is warmed from that temperature up to 150 °C. Work out its resistance when hot.

Increase in resistance
= $\alpha R_0 \Delta T$

= $4.3 \times 10^{-3} \times 10 \times 150$
= 6.45 Ω

Resistance when hot
= 10 + 6.45 = 16.45 Ω

Effect of length and diameter of resistor on resistance

Connect a light bulb (1 to 2 V) in series with a pencil lead (pencil 'lead' is made of carbon, an element frequently used in resistors) and a 6 V battery (four 1.5 V cells connected in series). Turn off the room lights and observe the brightness of the bulb.

Move one wire contact along the *length* of the pencil lead and observe the change in the intensity of light. The light should become brighter as the resistor becomes shorter, as there is less resistance impeding the circuit.

Repeat the activities with a pencil lead of different *diameter* and observe the changes in the intensity of light. The light should become brighter when the diameter of the resistor is larger as it is easier for the current to flow.

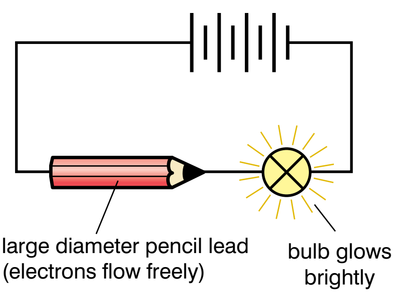


Figure 3.21 *A large diameter pencil lead, bright bulb*

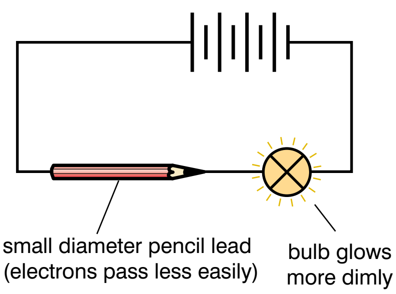


Figure 3.22 *A small diameter pencil lead, dim bulb*

The relationship between resistance R, resistivity ρ, length l and cross-sectional area A of a conductor

The resistance of a metal wire at a given temperature is determined by three factors:

- Its length l, in metres – the resistance is proportional to l, so if the length doubles so does the resistance.
- Its area of cross-section A, in m^2 – the resistance is inversely proportional to A, so a wire with twice the cross-sectional area will have only half the resistance.
- The material from which the wire is made – copper, for example, is a better conductor than iron.

Thus the resistance R of a wire can be expressed in the form:

$$R = \frac{\rho l}{A}$$

The symbol ρ (the Greek letter rho) is a constant, the value of which depends on the material from which the wire is made – the value for iron will be greater than the value for copper. We call ρ the resistivity of the material: it is defined by the equation above.

The units of resistivity are Ω m – ohms multiplied by metres. To see this, consider the units of the right-hand side. They will be Ω m × m divided by m^2, which works out correctly to give the resistance in ohms.

A tip for you

When you come to tackle a question on this topic, especially in an exam, you may have a moment of doubt: is the expression $R = \frac{\rho l}{A}$, or is it $R = \frac{\rho A}{l}$?

If you remember that the units of ρ are Ω m, you can quickly check that the second expression would give units for R of Ω m^2, which is wrong.

Worked example 3.9

What is the resistance of a copper cable that has a cross-sectional area of 1 cm^2 and a length of 2 km? The resistivity of copper is 2×10^{-8} Ω m.

Be careful over the units.

$l = 2$ km $= 2 \times 10^3$ m

$A = 1$ cm^2 $= 1 \times 10^{-4}$ m^2 (since there are 100 × 100 centimetre squares in a metre square)

$$R = \frac{\rho l}{A}$$

Putting in the values, we get

$$R = \frac{2 \times 10^{-8} \times 2 \times 10^{3}}{1 \times 10^{-4}} = 0.4\ \Omega$$

The relationship between resistivity and conductivity

The resistivities of most metals are in the range 10^{-7} to 10^{-8} Ω m. Those with larger resistivities conduct electricity less well.

The resistivity of an insulator such as dry polythene may be as high as 10^{15} Ω m.

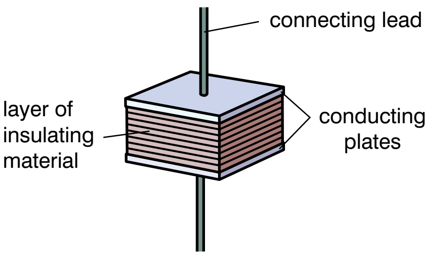


Figure 3.23 *The capacitor.*

Conductors and insulators are two extremes, but there are only a few materials in between these extremes: germanium at room temperature, for instance, may display a resistivity of around 0.001 Ω m. We call such materials **semiconductors**.

The capacitor (Figure 3.23) is an electrical component made of a combination of resistor and insulator. Many circuits use the capacitor to store and release charge. While there are many types of capacitor, they all have the same basic features – two conducting plates separated by a layer of insulating material (see page 71).

KEY WORDS

semiconductors *materials that have an electrical conductivity between that of a conductor and an insulator*

ammeter *an instrument used to measure the electric current in a circuit*

series circuit *an electrical circuit in which the current passes through each point in turn*

voltmeter *an instrument used to measure potential difference in a circuit*

potential difference *the difference in electrical charge between two points in a circuit measured in volts*

Constructing and drawing electric circuits

An electric current is a flow of charge. A circuit is made of conducting materials, with a device such as a cell, which keeps pumping the charges round the circuit.

Figure 3.22a shows the symbol for a cell. We often refer to it as a battery, though strictly speaking this is incorrect: ‘battery’ is the term reserved for a whole collection of cells acting together. The plus and minus signs are not usually marked on the drawing, so you must remember which is which. Remember too that **conventional current flows round the circuit from positive to negative**.

Figure 3.22b represents a battery of four cells. Notice that in order for them all to be pumping the charge the same way, the ‘+’ terminal of one cell has to be joined to the ‘–’ terminal of the next one.

Figure 3.24 *a) The symbol for a cell; b) a battery of four cells.*

The symbols for a cell, a light bulb, an ammeter and a switch are shown in the circuit in Figure 3.25. The **ammeter** is an instrument that measures the size of a current. It is fitted into the circuit so the current to be measured flows through it. In order that the ammeter does not reduce the current it has been put in to measure, it is essential that the instrument allows the current to flow through it freely (has a low resistance).

It is important that you connect the ammeter the correct way around, otherwise the current would try to twist it backwards. The terminal on it marked ‘+’ needs to be joined to the positive side of the battery, as shown in Figure 3.25.

(Sometimes the terminals are colour-coded. The ‘+’ terminal will be red, and the ‘–’ terminal black.)

Figure 3.26 shows the symbols for a range of electrical components. If you have access to equipment, test what you have learnt so far by using some of these components to make circuits such as that shown in Figure 3.25.

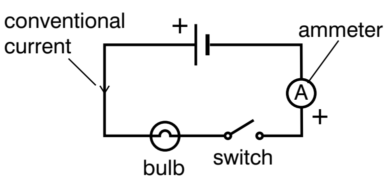


Figure 3.25 *A simple circuit.*

Try constructing the circuit shown in Figure 3.16 (page 93). Use the ammeter and **voltmeter** to make measurements, varying the current using the variable resistor. Record your measurements and discuss your findings with the class.

If you have access to other equipment – such as resistors, variable resistors, ammeter, voltmeter, capacitors – use these and observe what happens when you construct a circuit. Using the symbols for each component, draw a circuit diagram.

The correct position for an ammeter in a circuit

Do you think the current would be the same both before and after the light bulb in Figure 3.25? We will investigate this for the slightly more complex circuit of Figure 3.27 where there are two bulbs.

We call a circuit like this a **series circuit**: the current has to pass first through one bulb and then through the other. Notice again the way the three ammeters have been connected, with their '+' terminals nearer the positive side of the battery.

The result may surprise you:

the current has the same value all the way round a series circuit.

The correct position for a voltmeter in a circuit

Consider the circuit of Figure 3.28, which shows two resistors in series in a circuit. What is the voltmeter doing? You may think of the voltmeter as being connected in parallel (see Section 3.3 for a description of parallel circuits) with the first resistor. This is true, but it is probably not the best way of looking at it.

Unlike an ammeter, the voltmeter is not part of the circuit. If you wish to measure a voltage drop, you place the voltmeter on the bench nearby. One lead comes from the voltmeter to see what conditions are like at one point in the circuit, while the other lead goes to a second point in the circuit to find out what things are like there: the voltmeter then indicates the difference between the two points. We call it the **potential difference** (p.d.) between the ends of the resistor, and measure it in volts. Sometimes we speak of it as the voltage drop (or just the voltage) across the resistor.

The voltmeter shows the difference in something – but the difference in what? It might help you to think of a comparison with heat: a temperature difference makes thermal energy flow, and a potential difference makes charge flow.

Ideally a voltmeter should draw no current from the circuit it is investigating. It must therefore have a very high resistance.

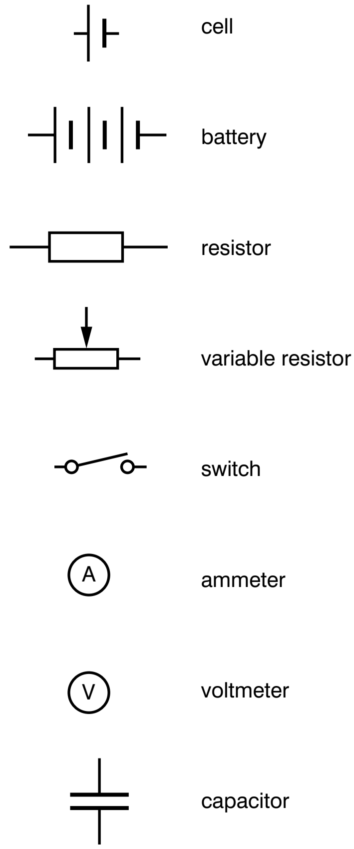


Figure 3.26 *Some circuit symbols.*

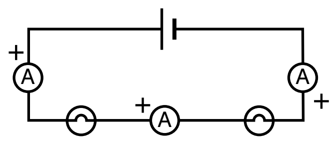


Figure 3.27 *Testing the current before and after it passes through the light bulbs.*

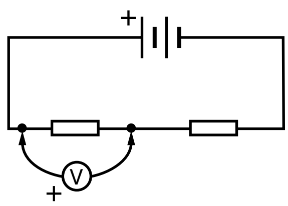


Figure 3.28 *A voltmeter connected in a circuit.*

Summary

- In a series circuit, the current is the same all the way round.
- You calculate current using $I = \frac{V}{R}$.
- You can measure the value of a resistor with an ammeter and voltmeter by applying Ohm's law.
- The resistance of a metal increases as its temperature rises. Its temperature coefficient of resistance (α) provides a numerical measurement of this.
- You can work out the resistance of a wire from $R = \frac{\rho l}{A}$, where ρ is the resistivity of the material of the wire.

Review questions

1. A battery of e.m.f. 6 V is connected to a circuit with a total resistance of 12 Ω. What current would you expect to flow?
2. A flashlight bulb has '2.4 V 0.3 A' marked on it. This means 'If you connect it to a 2.4 V battery a current of 0.3 A will be driven through it; this will make it white-hot'.
   a) Estimate the resistance of that flashlight bulb when it is lit up.
   b) If the bulb was connected to the mains (either 110 V or 220 V) instead, would the current initially be 0.3 A, more than 0.3 A or less than 0.3 A? Give your reason.
   c) What would happen next? Explain.
3. The resistivity of iron is 1.0×10^{-7} Ω m. Find the resistance of a 12 km length of a railway line with a cross-sectional area of 200 cm^2.
4. A tungsten resistor ($\alpha = 5.8 \times 10^{-3}\ K^{-1}$) with a value of 10.0 Ω at 0 °C is warmed from that temperature up to 150 °C. Work out its resistance when hot.
5. Figure 3.29 shows a circuit with a cell, a variable resistor and a lamp.
   a) Draw a circuit diagram of this arrangement.
   b) Redraw the circuit diagram adding an ammeter to measure the current in the circuit, and a voltmeter to measure the p.d. across the lamp.
   c) Describe how you would use this circuit to measure the current through the lamp for various p.d.s across it and describe the shape of a graph of p.d. against current that would be obtained.
   d) What would you change in this circuit to demonstrate Ohm's law?

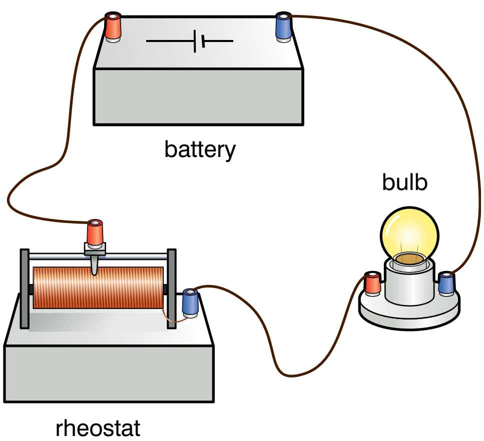


Figure 3.29

6. A student designs the circuit shown in Figure 3.30 in order to try to test Ohm's law.
   a) What is wrong with this circuit?
   b) Draw the circuit diagram the student should have used.
7. What is the difference between a cell and a battery?

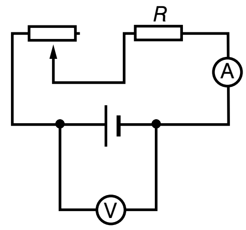


Figure 3.30

3.3 Combinations of resistors

By the end of this section you should be able to:

- Identify combinations of resistors in series, parallel and series–parallel connection.
- Derive an expression for the effective resistance of resistors connected in series.
- Derive an expression for the effective resistance of resistors connected in parallel.
- Calculate the effective resistance of resistors connected in series.
- Calculate the effective resistance of resistors connected in parallel.
- Calculate the current through each resistor in simple series, parallel and series–parallel combinations.
- Calculate the voltage drop across each resistor in simple series, parallel and series–parallel connections.

KEY WORDS

resistors *electrical components that limit or regulate the flow of electrical current in a circuit*

Resistors in different combinations

We shall now look at **resistors** connected in electric circuits in different ways – in series (Figure 3.31) and in parallel (Figure 3.34). In both cases, it might help to think of an electrical circuit as a large crowd of people trying to get into a stadium.

In Figure 3.29 overleaf, two resistors – 1 Ω and 3 Ω – are connected in series. Think of a line of people trying to get into a stadium, they have to queue up to get through one gate (1 ohm) and then through a narrower one (3 ohms) on their way into the stadium.

In Figure 3.35 on page 104 three resistors – 2 Ω, 3 Ω and 6 Ω – are connected in parallel. Think of this as the queue for the stadium again. In this case, the stadium has three (differently sized) gates for people to use, and they can enter the stadium much more easily.

Resistors connected in parallel are thus seen to have a lower total resistance than resistors connected in series.

In practice, electrical circuits can be quite complicated and can consist of resistors and other electrical components connected both in series and parallel.

Activity 3.6: Series and parallel circuits

Do these activities in groups.

Connect one bulb in series with a battery and note its brightness. Now connect a second and third bulb in series with the bulb.

Is there any change in the brightness of the first bulb when the second and third bulbs are added?

Once all bulbs are lit, remove one of the bulbs. What happens to the brightness of the others?

Connect one bulb to a battery and note its brightness. Now connect a second and third bulb in parallel with this bulb.

Is there any change in brightness of the first bulb when the second or third ones are added?

Once all bulbs are lit, remove one. Is there any change in the brightness of the remaining bulbs?

Series combination

Look at the circuit in Figure 3.31. There are two resistors (1 Ω and 3 Ω) in series. Each resistor offers a different resistance to the current flow and there is therefore a different voltage drop across each resistor.

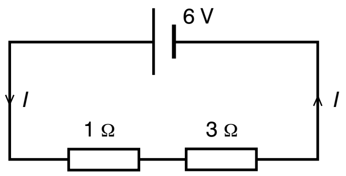


Figure 3.31 *Resistors in series in a circuit.*

With resistors in series, it is easy to work out the resistance of the circuit as a whole. You simply add up the separate values. In this case the circuit has a resistance of 4 Ω. (Be sure you could then work out that a current of 1.5 A will be drained from the battery, as $I = \frac{V}{R} = \frac{6}{4}$ from Ohm's law.)

For a series circuit, the total resistance $R = R_1 + R_2$

where R_1 and R_2 are the resistances of the separate components.

Worked example 3.10

Consider the circuit shown in Figure 3.32.

a) Calculate the total resistance of the whole circuit.

b) Calculate the current in the circuit.

Since the circuit is a series one, the same current must flow through all three resistors. Consider the 3.5 Ω resistor.

c) Calculate the voltage drop necessary to send 2 A through 3.5 Ω.

d) Calculate the p.d. across the other two resistors.

You will need to use the relationship $R = \frac{V}{I}$ in the form of $V = IR$ and $I = \frac{V}{R}$.

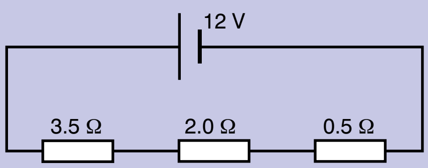


Figure 3.32

a) Total resistance in the circuit (series circuit)

$= 3.5 + 2 + 0.5 = 6\ \Omega$

b) Voltage supplied to the circuit is 12 V

Resistance in the circuit is 6 Ω

From Ohm's Law, $I = \frac{V}{R}$

$I = \frac{21}{6} = 2$ A

c) Resistance = 3.5 Ω

Current = 2 A

$V = IR$

$V = 2 \times 3.5 = 7$ V

The voltage drop across the 3.5 Ω resistor is 7 V.

d) For the 2 Ω resistor:

resistance = 2 Ω, current = 2 A

$V = IR = 2 \times 2 = 4$ V

For the 0.5 Ω resistor:

resistance = 0.5 Ω, current = 2 A

$V = IR = 0.5 \times 2 = 1$ V

Worked example 3.11

Calculate the size of resistor X in Figure 3.33.

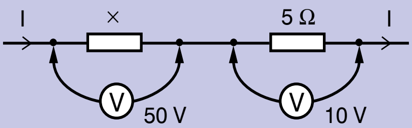


Figure 3.33

First look at the 5 Ω resistor. From $I = \frac{V}{R}$, a p.d. of 10 V across it must mean that a current I is flowing through it, given by:

$$I = \frac{V}{R} = \frac{10}{5} = 2 \text{ A}$$

Now look at resistor 'X'. Since they are in series the same 2 A is flowing through it. To achieve this there is a p.d. of 50 V across it.

$$R = \frac{V}{I} = \frac{50}{2} = 25\ \Omega$$

Parallel combination

It is less obvious how to work out the overall resistance of a circuit that consists of two resistors, not in series but in parallel (Figure 3.34).

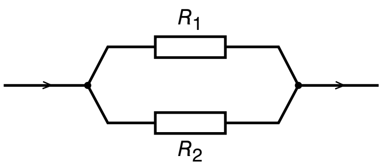


Figure 3.34 *Resistors in parallel.*

If the two resistors are R_1 and R_2, then they behave like a single resistor R the value of which is given by:

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$$

Worked example 3.12

A circuit consists of a 2 Ω resistor and a 3 Ω resistor in parallel. A 3 V battery is connected to the circuit. How big a current is drawn from it?

We must first calculate the effective resistance of the circuit using the formula for resistors in parallel.

Here,

$$\frac{1}{R} = \frac{1}{3} + \frac{1}{2} = \frac{3}{6} + \frac{2}{6} = \frac{5}{6}$$

That gives us 'one over *R*', so to find *R* we must turn it over.

$$R = \frac{6}{5} = 1.2\ \Omega$$

The question now becomes 'What current will a 3 V battery send round a 1.2 Ω circuit?'

$$I = \frac{V}{R} = \frac{3}{1.2} = 2.5\ \text{A}$$

The voltage drop across resistors in parallel

Look at the circuit of Figure 3.35. Remembering that there is no voltage drop down a conducting lead, you should be able to see that the left-hand end of all three resistors is at +6 V and their right-hand ends are at 0 V.

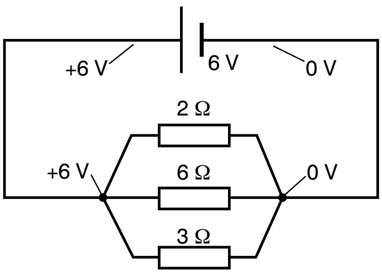


Figure 3.35

That illustrates what happens in this case:

if resistors are connected in parallel, they all have the same voltage drop across them.

This provides us with a means to work out the current in each branch of the circuit. Take the 2 Ω resistor as an example. The full 6 V of the battery is dropped across it, so the current *I* through it is given by:

$$I = \frac{V}{R} = \frac{6}{2} = 3\ \text{A}$$

Worked example 3.13

For the circuit shown in Figure 3.36:

a) calculate the voltage drop across each resistor

b) calculate the current through each resistor

c) calculate the effective resistance of the circuit.

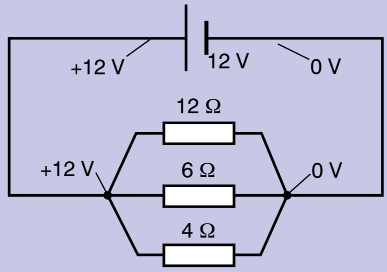


Figure 3.36

a) As the resistors are connected in parallel, the voltage drop will be the same across each, and will be equal to the full 12 V of the battery.

b) The voltage across the 12 Ω resistor is 12 V,

$I = \frac{V}{R} = \frac{12}{12} = 1$ A

The voltage across the 6 Ω resistor is 12 V,

$I = \frac{V}{R} = \frac{12}{6} = 2$ A

the voltage across the 4 Ω resistor is 12 V,

$I = \frac{V}{R} = \frac{12}{4} = 3$ A

c) Total resistance of the circuit is R.

$$\frac{1}{R} = \frac{1}{12} + \frac{1}{6} + \frac{1}{4}$$

$$= \frac{1+2+3}{12} = \frac{6}{12} = \frac{1}{2}$$

$R = 2\ \Omega$

Worked example 3.14

Two resistors of 10 Ω and 15 Ω are connected. What is their combined resistance if they are connected:

a) in series

b) in parallel?

a) In series

$R = R_1 + R_2$

$= 10 + 15 = 25\ \Omega$

b) In parallel

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$$

$$= \frac{1}{10} + \frac{1}{15}$$

$$= \frac{3+2}{30} = \frac{5}{30}$$

$$R = \frac{30}{50} = 6\ \Omega$$

Summary

- Resistors in series simply add.
- Resistors in parallel add by a '$\frac{I}{R}$' formula. Two resistors in parallel will conduct better than either one on its own.
- The resistance of the whole circuit includes the resistance within the battery itself, and this may not always be negligible.

- Where the circuit branches in a parallel circuit, the sum of the currents approaching the junction equals the sum of the currents leaving it.
- In a series circuit, the separate voltage drops (potential differences) across all the resistors add up to the voltage of the battery.
- In a parallel circuit each branch will have the same voltage drop across it.

Review questions

1. What should ammeters (a), (b) and (c) read in Figure 3.37?

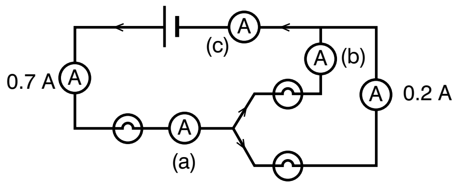


Figure 3.37

Which bulb should be the brightest? Why?

2. Give the missing ammeter readings a and b in Figure 3.38. Suggest why more current flows through some bulbs than through others.

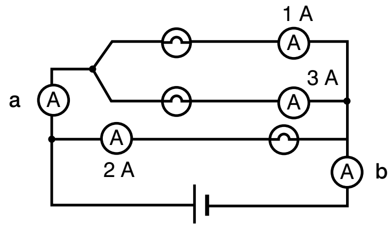


Figure 3.38

3. Give the missing ammeter readings c and d in Figure 3.39. Assume all the bulbs are identical ones. How would you expect the brightness of the different bulbs to compare?

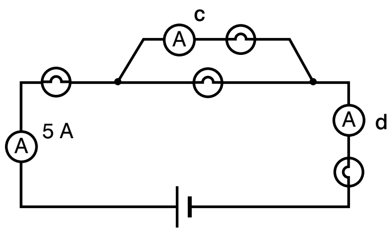


Figure 3.39

4. Draw a circuit diagram to show a battery of two cells with three light bulbs connected in parallel. Include two switches: one must turn one of the bulbs on and off, and the other must do the same for the other two bulbs together.
5. For the circuit diagram in Figure 3.40:
   a) What is the effective resistance of the two 4 Ω resistors in parallel?
   b) What is the total resistance of the circuit?
   c) Work out the current *I*.

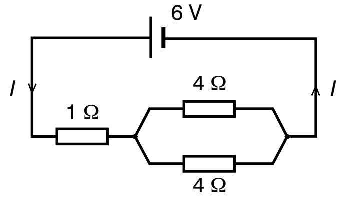


Figure 3.40

6. What single resistor could be placed in the box to replace the two shown in Figure 3.41?

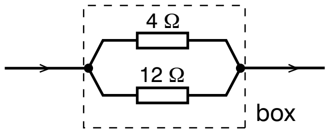


Figure 3.41

7. Calculate the effective resistance of a 25 Ω resistor and a 100 Ω resistor in parallel. What current will be drawn from a 50 V supply connected to the circuit?
8. a) Work out the total resistance of the circuit in Figure 3.42.
   b) What current will be drawn from the battery?

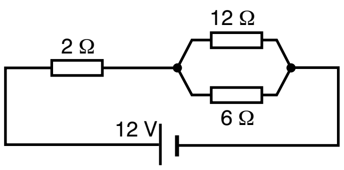


Figure 3.42

10. In the circuit of Figure 3.43 the battery has negligible internal resistance. Calculate:
    a) the total resistance in the circuit
    b) the current flowing in the circuit
    c) the potential difference across the 50 Ω resistor.

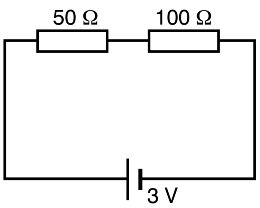


Figure 3.43

3.4 E.m.f. and internal resistance of a cell

By the end of this section you should be able to:

- Define the electromotive force (e.m.f.) of a cell.
- Distinguish between e.m.f. and terminal p.d. of a cell.
- Write the relationship between e.m.f., p.d., current and internal resistance in a circuit.
- Use the equation $V = E - Ir$ to solve problems in a circuit.
- Identify cell combinations in series and parallel.
- Compare the e.m.f. of combinations of cells in series and parallel.

The electromotive force of a cell

As we saw in Section 3.1, a cell pumps charge round a circuit. Its effectiveness at doing this is called its electromotive force (e.m.f.), in volts. A voltmeter is an instrument that can measure the e.m.f. of a cell: it has two leads going to it, and one is connected to each terminal of the cell.

Volts do not flow round a circuit. Volts cause coulombs of charge to flow at a rate of amperes. If we know the resistance R of a circuit, we can predict the current I that a battery will send round it, from $I = \frac{V}{R}$.

The resistance inside a cell

Suppose you short-circuit a cell. This means that you join its two terminals by a circuit that effectively has no resistance – a short piece of very thick copper wire, for instance. The cell has an e.m.f. V, but the circuit apparently has no resistance R. What happens then? Does the current increase without limit?

The point we are forgetting is that the cell has to pump the charge round the whole circuit, and that includes the part within the cell as well as the external circuit. Internal resistance varies between different cells, but it is what finally sets a limit to the current a cell can supply.

A 1.5 V torch cell (represented in Figure 3.45 by the dotted line) typically has an internal resistance of up to an ohm. This means that even if you short-circuit the cell, there is still an ohm of resistance. The biggest current it can deliver is given by $I = \frac{V}{R} = \frac{1.5}{1} = 1.5$ A.

So far, we have taken it for granted that if you double the e.m.f. in a circuit, you will double the current. This should be checked experimentally, but it needs some careful thought in order to give it a fair test.

The obvious thing to do is to add cells one at a time to a circuit, measure the current and see if it increases accordingly. A fair test

Activity 3.7: Measuring electromotive force and terminal voltage of a cell

Set up a circuit consisting of a dry cell, a bulb, a switch and a voltmeter. Connect the voltmeter as shown in Figure 3.44.

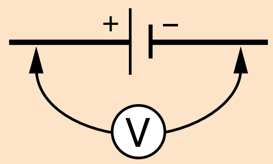


Figure 3.44

Take the reading in the voltmeter while the switch is on.

Then another reading while the switch is off.

Compare the two readings.

Which one is larger? Which reading is the e.m.f.? Discuss with your group.

requires that you change only one thing at a time, which means that you must not change the resistance of the circuit.

There are two problems here:

1. Adding another cell adds a bit more resistance to the circuit. All you can do about that is to choose cells that have a very low internal resistance (such as the lead–acid cells that make up a car battery) or to use a circuit for which the resistance is so high that another ohm or two makes virtually no difference.
2. The resistance of a metal wire changes as it heats up. This rules out circuits containing light bulbs, because as you add more batteries their temperature changes from less than red-hot to strongly white-hot: a change of perhaps 1000 degrees!

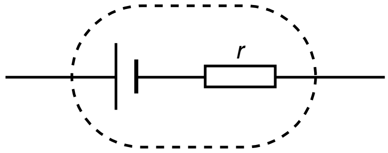


Figure 3.45 *Representation of a cell indicating its internal resistance r.*

The difference between e.m.f. and terminal p.d. of a cell

A high resistance voltmeter connected across the terminals of a cell can give a good reading of the cell's e.m.f. However, if the cell is then connected in a circuit where the current is high, the reading on the voltmeter drops. This drop is caused by the cell's internal resistance.

Figure 3.46 shows the cell from Figure 3.45 connected in a circuit.

If the e.m.f. of this cell is E and it is connected to a resistor R, the potential difference across its terminals (V) will now be less than E because some of its energy is used to drive the current (I) through the internal resistance (r) in the cell. If R decreases, the current (I) increases and the terminal p.d. (V) of the cell will decrease.

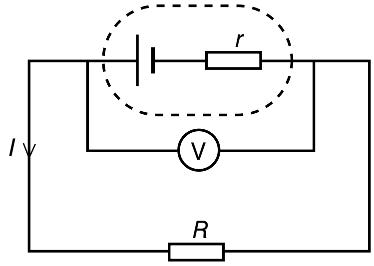


Figure 3.46

The relationship between e.m.f., current and internal resistance in a circuit

The relationship between a cell's e.m.f. (E) and internal resistance (r) and the current (I) and resistance (R) in a circuit is given by the expression:

E = p.d. across R + p.d. across r

$E = IR + Ir$

$E = I(R + r)$

As $V = IR$ (Ohm's law) this can also be written as $E = V + Ir$, or $V = E - Ir$.

Worked example 3.15

In the circuit shown in Figure 3.46, the cell's e.m.f. is 10 V and its internal resistance is 2 Ω. Find the p.d. across the terminals of the cell when it is connected to a 3 Ω resistor.

First find the current in the circuit:

$E = I(R + r)$

$$I = \frac{E}{(R + r)}$$

$$= \frac{10}{(3 + 2)} = 2 \text{ A}$$

The p.d. across the cell terminals is equal to the p.d. across the resistor (3 Ω). Find the p.d. across the resistor:

$V = IR = 2 \times 3 = 6$ V

Therefore the p.d. across the terminals of the cell is 6 V.

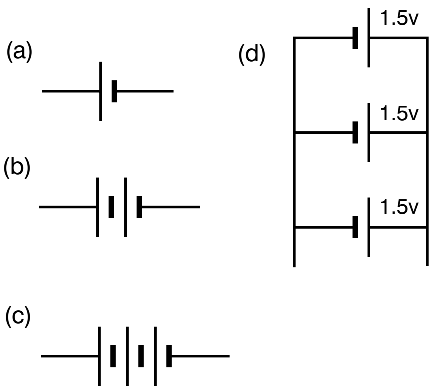


Figure 3.47 Combining cells.

Combinations of cells in series and parallel

In Figure 3.47, (a), (b) and (c) show cells connected in series, (d) shows cells connected in parallel. Each cell has an e.m.f. of 1.5 V.

The total e.m.f. of cells connected in series is the sum of the e.m.f. of each cell: $E = E_1 + E_2 + \ldots + E_n$. Equipment requiring high power uses cells arranged in series.

The e.m.f. of cells connected in parallel is the e.m.f. of an individual cell: $E = E_1 = E_2 = \ldots = E_n$. The current supplied by a parallel arrangement of cells can be maintained for far longer than that supplied by a single cell, so equipment requiring a steady current for long periods will use cells arranged in parallel.

(a) 1 cell	total e.m.f. = 1.5 V
(b) 2 cells in series	total e.m.f. = 1.5 + 1.5 = 3 V
(c) 3 cells in series	total e.m.f. = 1.5 + 1.5 + 1.5 = 4.5 V
(d) 3 cells connected in parallel	total e.m.f. = 1.5 V

Activity 3.8: Connecting cells together

For this activity you can use either the fruit cells made in Activity 3.3 or purchased 1.5 V cells. Connect a small light bulb in a circuit with one cell. Observe its brightness. Now take another cell and connect it in the circuit as shown in Figure 3.48a. Observe the brightness of the bulb. You can add further cells in this way if you wish.

The cells in Figure 3.48a are connected in series. Now connect two cells together in parallel, as shown in Figure 3.48b. Now connect these cells to a light bulb as shown in Figure 3.48c. Observe its brightness.

How does the brightness of the bulb in the circuit with one cell compare with its brightness in the circuit with two cells in series?

How does the brightness of the bulb in the circuit with the cells in parallel compare to the brightness when the cells are in series?

You can see that the cells are connected negative to negative, positive to positive in Figure 3.48b. See what happens if you connect them negative to positive and positive to negative before connecting them in a circuit. What do you find?

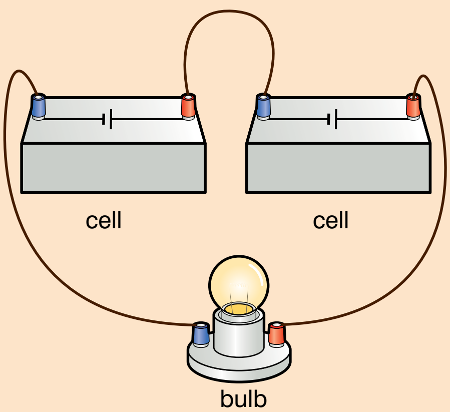


Figure 3.48a

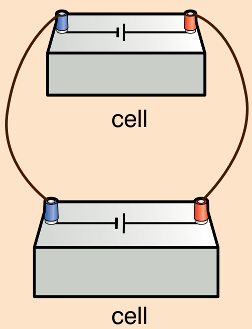


Figure 3.48b

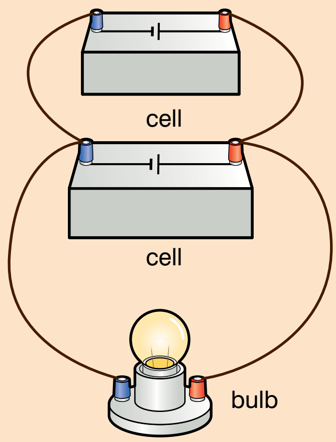


Figure 3.48c

Review questions

1. Imagine you have a brother two years younger than you are. Explain to him why:

   a) the '+' terminal of one cell in a circuit should be connected to the '–' terminal of a second cell, but

   b) the '+' terminal of an ammeter needs to be connected to the '+' terminal of the battery.

2. A circuit consists of three resistors in series: one of 3 Ω, one of 7 Ω and one of 10 Ω. A battery of e.m.f. 12 V and negligible internal resistance is connected to the circuit. What size current will it supply?

3. A battery has an e.m.f. of 3 V and an internal resistance of 1 Ω. What current will it give if it is connected to a circuit of resistance:

   a) 2 Ω

   b) 4 Ω?

   What current will it give if its terminals are short-circuited?

4. In Figure 3.49, A and B represent the terminals of a battery of e.m.f. 4 V and internal resistance 0.5 Ω. *R* is the total resistance of the circuit to which it is connected.

   a) Explain why a voltmeter connected as shown across *R* would read the same as a voltmeter connected across the terminals of the battery.

   b) Calculate the current which flows round the circuit if the resistance of *R* is 1.5 Ω.

   c) What then is the p.d. between the terminals of the battery?

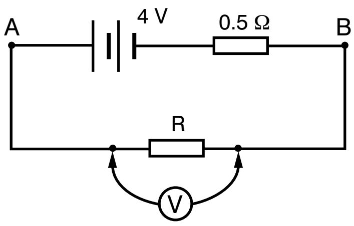


Figure 3.49

5. a) What current is flowing through the 3 Ω resistor in Figure 3.50?

   b) If its internal resistence is 2 Ω , what is the e.m.f. of the battery?

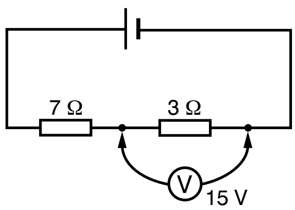


Figure 3.50

3.5 Electric energy and power

By the end of this section you should be able to:

- Define electrical energy and power in an electrical circuit.
- Find the relationship between KWh and joule.
- Use $P = VI = \frac{V^2}{R} = I^2R$ to solve problems in electric circuits.
- Use $W = VIt = I^2Rt = \frac{V^2t}{R}$ to calculate electric energy dissipated in an electric circuit.
- Calculate the cost of electrical energy expressed in KWh.

Worked example 3.16

A 60 W light bulb is switched on for 2 minutes. How many joules of electrical energy does it convert into heat and light in that time?

The power of the light bulb is 60 W; this means that it uses 60 J of energy every second.

It is switched on for 2 minutes = 120 s.

Energy = VIt (and Power = VI)

Total energy supplied = 60 J/s × 120 s = 7200 J

Definitions for electrical energy and power

The word 'power' here means the rate at which energy is being supplied or converted. Power is measured in the units of joules per second (J/s), or watts (W).

The electrical energy produced by a current of I amperes flowing through a p.d. of V volts for a time t seconds is given by:

Energy = VIt J.

Power is the energy produced by an electrical appliance in one second. Thus, if $t = 1$:

Power = $VI \times 1 = VI$ W.

Worked example 3.17

If the potential difference across a working electrical motor is 50 V and the current is 2 A, calculate the power of the motor.

Power = $VI = 50 \times 2 = 100$ W

Consider a resistor that converts all the electrical energy supplied to heat.

The power produced = VI

This can be rewritten (using Ohm's law, $V = IR$) as $IR \times I = I^2R$

Therefore, for a resistor

heat produced per second = power = I^2R.

If V and R are known, then the equation can be written (again using Ohm's law)

power = $VI = V \times \frac{V}{R} = \frac{V^2}{R}$.

Worked example 3.18

Two heating coils A and B produce heat at a rate of 1 kW and 2 kW, respectively, when connected to 250 V mains.

a) Calculate the resistance of each resistor.

b) Find the power they would produce when connected in series to the mains.

a) For the first resistor:

Power = VI

$1000 = 250 \times I$, so $I = \frac{1000}{250} = 4$ A.

So $R = \frac{V}{I} = \frac{250}{4} = 62.5\ \Omega$

For the second resistor:

Power = VI

$2000 = 250 \times I$, so $I = \frac{2000}{250} = 8$ A.

So $R = \frac{V}{I} = \frac{250}{8} = 31.25\ \Omega$

b) If the resistors are wired in series, their total resistance is

$R_1 + R_2 = 62.5 + 31.25 = 93.75\ \Omega$.

using $R = \frac{V}{I}$

$93.75 = \frac{250}{I}$, giving $I = \frac{250}{93.75}$

$= 2.67$ A.

So power = $VI = 250 \times 2.67 = 666$ W

Notice that this is less than the power produced when either resistor is connected to the mains separately. Can you see why?

Worked example 3.19

Calculate the power of a water heater that draws a current of 10 A from a 220 V supply.

Power = $VI = 10 \times 220 = 2.2$ kW

Cost of electrical energy

Electricity is distributed to homes and businesses and the quantity supplied is measured in kilowatt hours (KWh). A kilowatt hour is the energy used by a 1 kW appliance working for 1 hour. Consumers of electricity are charged for each kWh used.

How many joules are there in 1 kWh?

One kWh = energy transformed by 1 kW (1000 W) for 1 hour (3600 s) = energy when 1000 J are transformed each second for 3600 s = 1000 × 3600 = 3 600 000 J = 3.6 MJ

Worked example 3.20

How many units (kWh) are used by:

a) a 3 kW electric fire used for 2 hours,

b) a 100 W light bulb used for 15 hours?

a) Number of units = $3 \times 2 = 6$ kWh

b) Number of units = $0.1 \times 15 = 1.5$ kWh

Activity 3.9 Power requirements

Which appliances in your home consume the greatest amount of energy? Is it the refrigerator? The TV set? The electric stove? You can find out by inspecting appliances in your home and determining the number of watts each consumes and multiplying this by the number of hours operated. By law, each electric device must specify power requirements, and these are generally recorded on a small tag located on the appliance or on the power cable connected to it. Inspect all the appliances in your home and record power requirements in a table.

Worked example 3.21

A 1.5 kW electric fire is accidentally left on overnight for eight hours. The cost of a unit of electrical energy is 0.273 Birrs. How much money has been wasted?

The number of units (kWh) = power in kW $\times$ time in hours =
1.5 $\times$ 8 = 12 kWh

The total cost = 12 $\times$ cost of each unit = 12 $\times$ 0.273 = 3.28 Birrs

Worked example 3.22

A current of 4 A flows through an electric fire for 1 hour. The supply voltage is 240 V. What energy is transformed by the fire in 1 hour? (1 hour = 60 $\times$ 60 seconds)

$W = VIt = 240 \times 4 \times 3600 = 3.456$ megajoules (MJ)

Worked example 3.23

The Ethiopian Power Corporation is distributing power saving lamps to the public free of cost. If an 11 W (0.011 kW) power saving bulb is used in place of the equivalent 60 W (0.06 kW) conventional lamp, and each type of bulb is used for 10 hours a day for four weeks (seven days a week), how much would the customer of the Ethiopian Power Corporation save?
Number of hours = 10 $\times$ 7 $\times$ 4 = 280
Cost of one unit of electricity is 0.273 Birrs

Conventional lamp
Number of units (kWh) = power in kW $\times$ time in hours = 0.06 $\times$ 280 = 16.8 kWh
Total cost of using conventional bulb = 16.8 $\times$ 0.273 = 4.59 Birrs

Power-saving lamp
Number of units = 0.011 $\times$ 280 = 3.08 kWh
Total cost of using power-saving lamp = 3.08 $\times$ 0.273
= 0.84 Birrs

The customer would save 3.75 Birrs.

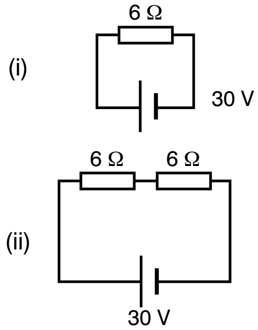


Figure 3.51

Review questions

1. Work out the sources of energy used at home. Can you suggest any economies or improvements?
2. Carry out the same analysis for your school.
3. In the circuit of Figure 3.51(i) calculate:
   a) the current flowing
   b) the number of joules per second (power) of heat produced in the 6 Ω resistor.

Another 6 Ω resistor is added in series with the first (Figure 3.51(ii)). Calculate:

c) the current flowing now

d) the power produced in the two 6 Ω resistors together

e) the power produced in a single 6 Ω resistor.

Comment on the reasons for the difference in the heat produced in a 6 Ω resistor in the two circuits.

3.6 Electric installation and safety rules

By the end of this section you should be able to:

- Understand the dangers of mains electricity.
- Have some awareness of safety features incorporated in mains electrical installations.
- Understand the nature of the generation and supply of electricity in Ethiopia.
- Consider employment prospects in Ethiopia's electricity industry.

KEY WORDS

fuse *protective device for protecting an electrical circuit, containing a wire that melts and breaks when the current exceeds a certain value*

In Ethiopia, while traditional sources of energy, such as wood, are still of great importance, the provision of mains electricity is becoming increasingly significant. Mains electricity is generated, transmitted and distributed by the Ethiopian Electric Power Corporation (EEPCO).

While Ethiopia uses some fossil fuels for electricity generation, the country is fortunate in having access to sources of hydroelectric power at Melka Wakena, Finchaa, Koka, Awash, Tis Abay, Gilgel Gibe and other planned sites, for electricity generation. Ethiopia also has potential for using other renewable energy sources, such as solar, wind and geothermal energy, for generating electricity.

Work is underway to increase the availability of mains electricity in Ethiopia. Mains electricity is much more powerful than the electricity we have studied in this unit, and great care must be taken when using it. Any work being undertaken on mains electrical installations must only be undertaken by those who have qualified through the Electrical Professional Competence Certification Scheme (operated by the Ethiopian Electricity Agency).

Various safety features are incorporated into mains electrical installations; one of these features is the **fuse**, which uses the heating effect studied earlier in this unit. A fuse is a small, thin piece of wire in a glass tube that has metal connectors on each end. It is designed to be used in a circuit to allow a current of a certain size to pass through it. If the current in the circuit increases beyond this point, the fuse heats up and then breaks, causing the current in the circuit to stop. The fuse is a very useful safety device, protecting appliances from surges of current.

KEY WORDS

earthed *circuit connected to the earth, allowing any dangerously large current to be safely discharged*

Electrical safety is also increased if appliances are **earthed**. This means that if there were a fault in an appliance, any dangerous large current would run harmlessly to the earth and cause the fuse to blow, thus preventing injury.

Miniature circuit breakers (MCBs) can now be used in place of fuses in some appliances. They cut off the power if there is a fault that causes the appliance to overheat. They can be re-set when the appliance cools down.

Activity 3.10: Electrical safety

Take five 1.5 V cells, a 2 A fuse and a selection of connecting wires. Connect the circuit shown in Figure 3.52.

Look carefully at the fuse, and while observing the fuse closely, close the switch.

The fuse consists of a very fine length of wire inside a transparent tube with metal connectors at each end.

When the switch is closed, the wire in the tube quickly becomes red and then separates into two pieces. The wire becomes dark again as it cools down. At the end of the experiment, the wire in the fuse has a gap in the centre.

There might be a small lump of metal on one of the broken ends of wire. This is a result of the metal becoming very hot and melting as the current passes through it, before the fuse wire breaks and causes the current to stop.

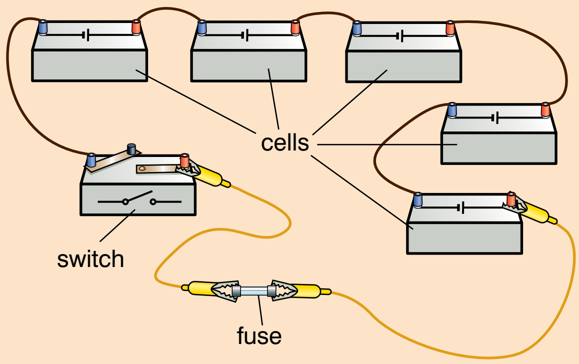


Figure 3.52

Engineering project

New designs for cars employ hybrid motors. Cars use both electricity and gasoline for power. When going down hill, instead of standard brakes, which convert kinetic energy into heat, these cars capture energy by using 'electric brakes' that recharge the battery or convert the energy into rotational energy in a massive fly wheel.

Try designing on paper a hybrid car that captures as much energy as possible. In teams you may wish to build a model of your design, showing innovative ways to save energy.

Earth circuit leakage breakers (ECLBs) are used in place of fuses in domestic electricity supply boards, where they are useful in cutting off the supply very quickly if a fault occurs.

Small, plug-sized ECLBs (also called residual current circuit breakers) are also useful in preventing injury when using electrical equipment outdoors. They will cut off the power supply very quickly if the power cable is cut accidentally.

Review questions

1. For safety, the circuit for a 3 kW metal-bodied electric kettle is both earthed and contains a fuse.
   a) Why is the kettle earthed?
   b) Why does the circuit contain a fuse?
2. In a house you might find i) MCBs, ii) an ECLB.
   a) Say where each of these may be found.
   b) Explain the purpose that each might be serving.

End of unit questions

1. If 36 C of charge pass through a wire in 4 s, what current is it carrying?
2. An appliance has a resistance of 5 Ω and requires a current of 0.5 A for it to work. The only battery available has an e.m.f. of 12 V and negligible internal resistance. What extra resistance will you need to provide to limit the current to its correct value. Draw the circuit you would assemble.
3. A student incorrectly set up the circuit shown in Figure 3.53.

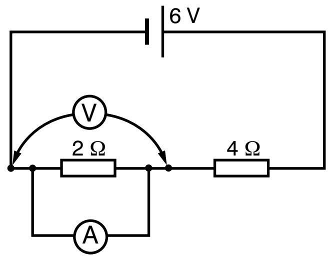


Figure 3.53

   a) What is wrong with it?
   b) Assume that both the ammeter and the battery have negligible resistance. Starting from one terminal of the battery, what route will all the current take as it flows round the circuit? How much resistance will this circuit have? What will the ammeter read?
   c) What will the voltmeter read?
   d) The student then reconnects the circuit so that the two resistors, the ammeter and the voltmeter are all in series with the battery. What would you expect the ammeter to read now? Give a reason for your answer.
4. A 12 V light bulb is put under test, with the following results:

Current through bulb (A)	0.4	0.6	0.8	1.0	1.2	1.4	1.6
p.d. across bulb (V)	1.4	2.6	3.9	5.5	7.4	9.7	12.6

   a) Draw the circuit you would use to obtain these readings.
   b) Copy the table out, and for each set of readings work out the resistance R of the bulb.
   c) Plot a graph of R against I. Why is it not valid to assume the graph goes through the origin?
   d) Use your graph to estimate the resistance of the bulb at room temperature. (Hint: in what circumstance will the bulb not be heated at all?)
5. Constantan is the name of an alloy that is sometimes used for making resistors in the laboratory. Its resistivity is 4.9×10^{-7} Ω m. Calculate the resistance of a 3 m long constantan wire with 1 mm^2 cross-sectional area.
6. Draw a labelled diagram of a dry cell. State the function of each of its parts.
7. The resistivity of copper is 2×10^{-8} Ω m. Work out the resistance of a copper wire of 1 mm^2 cross-sectional area and 3 m long.

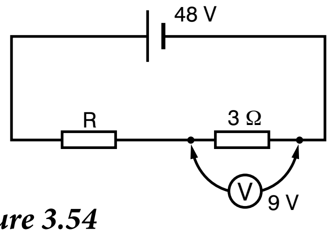


Figure 3.54

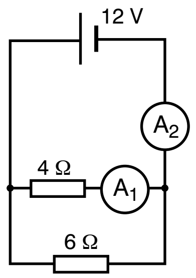


Figure 3.55

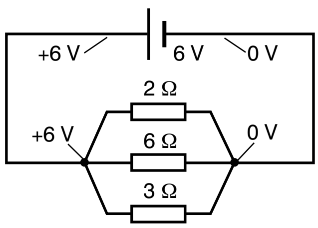


Figure 3.56

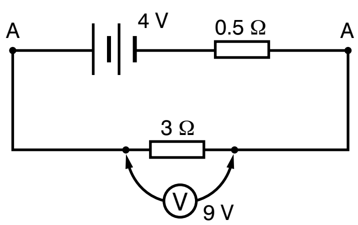


Figure 3.57

8. A battery sends a current of 3 A through a 4 Ω resistor.
   a) What must the e.m.f. of the battery be?
   b) How many coulombs of charge will pass through the resistor in 1 minute?
9. Explain why it is wrong to write the units of resistivity as ohm/metre (Ω/m) with the stroke between the two quantities.
10. Draw a circuit diagram to show a battery of two cells and three light bulbs connected in series. Include a switch to turn the lights on and off. Does it matter where the switch is placed in the circuit?
11. a) What current is flowing round the circuit in Figure 3.54?
    b) How great must the p.d. be across resistor R?
    c) What is the resistance of resistor R?
12. What should be the reading of ammeters A_1 and A_2 in Figure 3.55?
13. These questions refer to the circuit drawn in Figure 3.56.
    a) What are the currents through the 6 Ω branch, through the 3 Ω branch and through the 2 Ω branch?
    b) What is the total current drawn from the battery?
    c) As far as the battery is concerned, what is the total resistance of its whole circuit?
    d) Compare this with the value obtained from

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

14. In Figure 3.57, A and B represent the terminals of a battery of e.m.f. 4 V and internal resistance 0.5 Ω. *R* is the total resistance of the circuit to which it is connected. Calculate the current that flows round the circuit and the p.d. between the terminals of the battery if the resistance *R* is
    a) 7.5 Ω
    b) 0 Ω.

Electromagnetism

Unit 4

Contents

Contents

Section	Learning competencies
	• State Lenz's law. • Indicate the direction of induced currents, given the direction of motion of the conductor and the direction of a magnetic field. • Describe the factors that affect the magnitude of induced e.m.f. in a conductor. • Describe the link between electricity and magnetism. • Apply Faraday's law to calculate the magnitude of induced e.m.f. • Define inductance and its SI unit. • Distinguish between self- and mutual inductance. • Apply the definition of inductance to solve simple numerical problems. • Explain the action of the simple a.c. generator. • Compare the actions of d.c. and a.c. generators. • Draw a diagram of a transformer. • Give a simple explanation of the principles on which a transformer operates. • Identify that, for an ideal transformer, $P_{out} = P_{in}$. • Show that, for an ideal transformer, $V_s/V_p = N_s/N_p$. • Apply the transformer formulae to solve simple problems.

KEY WORDS

magnetic force *the force exerted between magnetic poles, producing magnetisation*

domains *regions within a magnetic material which have uniform magnetisation*

4.1 Magnetism

By the end of this section you should be able to:

- Investigate the domain theory of magnetism.

The force between magnets

In Grade 9 you learnt about the **magnetic force**. If you bring two bar magnets towards each other, then you will feel either a force of attraction between them or a force of repulsion.

Activity 4.1: The force between two bar magnets

Use two bar magnets to remind yourself about the force between them. Set them up as shown in the Figure 4.1.

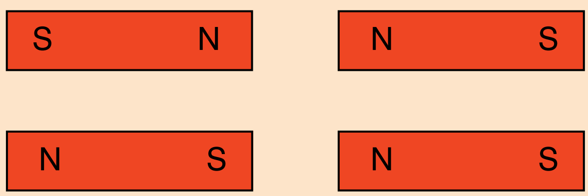


Figure 4.1 *Force of attraction or repulsion between two bar magnets.*

The Earth's magnetic field

The results in Activity 4.1 can be explained by saying that each bar magnet has two 'poles', which we call 'north seeking' and 'south seeking'. (These are often marked with an S or an N on the magnet.) Like poles will repel each other but unlike poles will attract. What other things have you met in physics that behave in a similar way? (Look back to Unit 2 if you cannot answer this question!)

The word 'pole' may seem a strange term to use for the ends of a magnet. However, it is used because if a bar magnet is suspended, its poles will point in the direction of the north and south poles of the Earth. (They 'seek' the Earth's poles.) This happens because the Earth is like a huge magnet and it has a magnetic field, which acts between the poles. You will learn more about magnetic fields in Section 4.2, but for the moment you simply need to know that the Earth has a magnetic field and this is what enables compasses to show you which way is north.

A compass needle is like a small suspended bar magnet. Its north-seeking pole will point to the Earth's north pole. If you know which direction is north, then you can line up a map correctly to find your way through unfamiliar areas.

What are magnetic domains?

Imagine a piece of steel to be made up of a huge number of invisibly small magnets. This does not explain what magnetism is, of course; it just replaces the question 'What is a magnet?' by the question 'What are those tiny magnets?' Nevertheless it does help us to make some progress, as it explains several aspects of the behaviour of complete magnets.

Let us start with an ordinary unmagnetised piece of steel. The little magnets are there, but they are in clusters rather like you would get if you threw a large number of small bar magnets into a box. The 'N' end of one magnet is up against the 'S' end of another, and the two effectively cancel each other out (see Figure 4.2a). We call these magnetic **domains**.

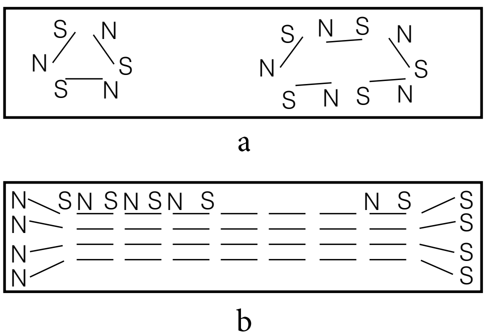


Figure 4.2 *Magnetic domains.*

Figure 4.2b shows the steel when it is fully magnetised. Notice that there are a large number of north poles at one end of the bar, and the same number of south poles at the other. On this picture we would not expect to find a magnet that had a north pole unless

DID YOU KNOW?

Birds can detect the Earth's magnetic field, and use the field to find their way when they migrate from cold to warm climates or from warm to cold climates. Cows stand or lie so that their bodies are in a north–south direction but the magnetic fields surrounding high voltage power lines can confuse their sense of direction!

Activity 4.2: Using a compass

In a small group, go outside and use a compass to find out which direction is north. Use this information to chalk directions on the ground: N, NE, E, SE, S, SW, W, NW. Using these directions, and non-standard measurements such as paces, one member of the group should give the others directions to reach particular objects, such as trees. Repeat this in several locations in your school compound.

it also had a corresponding south pole somewhere else. A single magnetic pole, all on its own, has never been found, although from time to time scientists have looked for them.

Activity 4.3: To investigate magnetic domains, magnetisation and demagnetisation

Fill a test tube two-thirds full with iron filings or shredded steel wool and bring the end of the tube towards first the north end of a compass needle and then the south end, as shown in Figure 4.3.

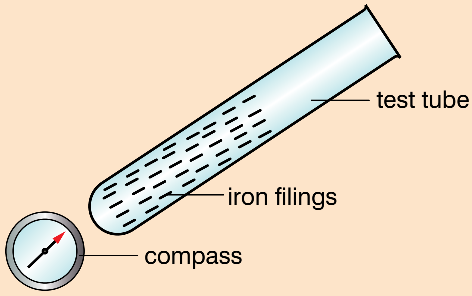


Figure 4.3 *Procedure for the investigation.*

Is one end of the needle more attracted to the tube than the other? Record the maximum angle to which the needle is deflected (see Figure 4.4).

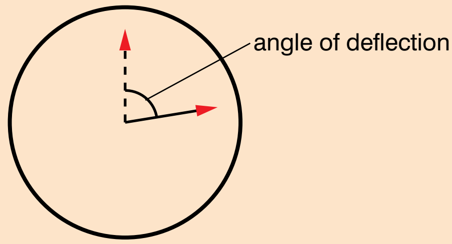


Figure 4.4 *Angle that needs to be recorded.*

Now stroke the tube 50 times with a permanent magnet and repeat the procedure. Record the results.

Finally, shake the tube vigorously for one minute (make sure the tube is firmly sealed otherwise the iron filings will go everywhere!). Record the results.

In which situation was the test tube most highly magnetised (when was the angle of deflection of the compass needle greatest)? Why do you think this happened? Try to explain the results to a partner, in terms of magnetic domains, before reading on.

KEY WORDS

magnetisation *the extent or degree to which an object is magnetised*

Explaining the results

In the first situation, the magnetic domains in the tube were arranged randomly so the end of the tube was not strongly magnetised. When the tube was stroked by a permanent magnet, the magnetic domains inside the tube arranged themselves so that they were lined up. The end of the tube was magnetised and so the compass needle deflected to a greater angle. Shaking the tube vigorously disturbed this arrangement and demagnetised the end of the tube.

If the steel is weakly magnetised, this means that some of its domains are lined up but some are still in clusters. This picture is supported by experiments that show there is a limit to how strongly a given bar can be magnetised ('magnetic saturation' has been reached). This occurs when all those tiny magnets are in lines and there are no more domains left to break up.

Magnetisation

Activity 4.3 showed that once you have magnetised your steel to give a magnet, it may not last in that state for ever. Figure 4.5 is a drawing of one end of a magnetised piece of steel. It is the repulsion between like poles that is forcing the tiny magnets to spread out there, and the whole situation is rather unstable.

Imagine holding a set of bar magnets all the same way round side by side in your hand. There will be a very marked tendency for some of them to turn right round so that opposite poles come together. In a similar way, it will not take much for those tiny magnets repelling each other at the end of the steel to jump round and form a cluster again. In other words, a magnet is liable to become weaker as time passes.

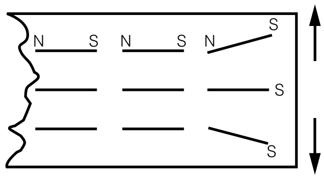


Figure 4.5 *One end of a magnetised piece of steel.*

Activity 4.4: Magnetisation by heating and cooling

Use insulated tongs or pliers to heat a nail in the hottest part of the flame of a Bunsen burner until it glows. Use the tongs or pliers to remove the nail from the flame and place it lengthwise on a permanent bar magnet (see Figure 4.6a). Record how long the nail is in contact with the permanent magnet. After the nail has cooled, measure its magnetic strength by finding out how many paper clips can be suspended from a chain at one end (see Figure 4.6b).

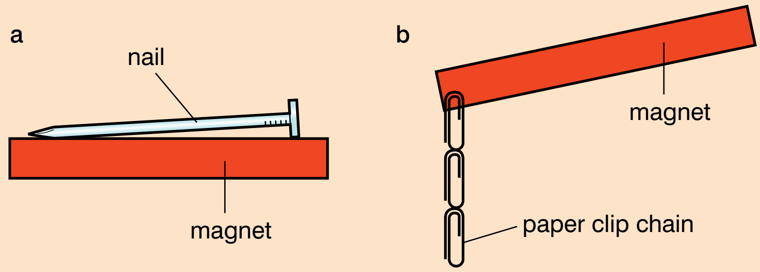


Figure 4.6 *Set up for the investigation.*

Use another nail that has not been heated and put in on the same bar magnet for the same length of time. Measure the magnetic strength of this nail by recording how many paper clips can be suspended from a chain at one end. Discuss your results in a small group. Try to explain what happened in terms of the magnetic domains before you read on.

Explaining the results in terms of magnetic domains

The tiny magnets, or domains, of the nail are the individual atoms that make up the nail, though what makes them behave in that way is a question we cannot answer at this stage. These atoms will be moving slightly from a given position, and as the nail is heated they will move more and more. You might expect that this movement is likely to encourage them to jump round out of line (think of a box containing a lot of bar magnets all nicely lined up, and imagine giving it a good shake). By heating the nail until it is red hot, you cause all the domains to jumble up. As the nail cools on a permanent magnet, the jumbled up domains line up again and the nail becomes magnetised once more.

Activity 4.5: Magnetic shielding

Support a bar magnet using a stand on a table. Attach a thread to a paper clip and hang it below the magnet (see Figure 4.7).

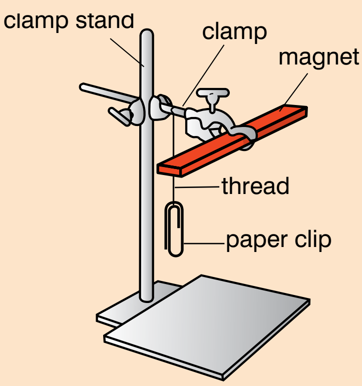


Figure 4.7 *Set up for the activity.*

Record what you see. Now slide a small sheet of paper in the gap between the paper clip and the magnet, being careful not to touch either of them! Record what you see. Repeat this using a sheet of plastic, a piece of aluminium foil and the lid of a tin can. Organise your observations in a table. Try to explain your results in terms of magnetic shielding.

KEY WORDS

magnetic shielding *limiting the penetration of magnetic fields using a barrier of conductive material*

Some of the domains in the cold nail will be jumbled up as well, but some will be lined up before the nail is placed on the permanent magnet. As with the nail that has been heated, the permanent magnet will cause the rest of the domains to line up so that the nail becomes magnetised.

Magnetic shielding

Sometimes you may not want a piece of equipment to become magnetised as it could be damaged as a result. A **magnetic shield** stops the equipment from being affected by a magnet. Magnetic shielding comes in various forms depending on the equipment that needs to be protected.

Activity 4.6: Which material makes the best magnetic shield?

Cut the bottoms from two paper cups of different sizes, two plastic cups of different sizes and two tin cans of different sizes.

Place a compass on a table and record the direction of magnetic north. Now place two bar magnets 7 cm to the east and west of the compass so that the north pole of one faces the south pole of the other, as shown in Figure 4.8.

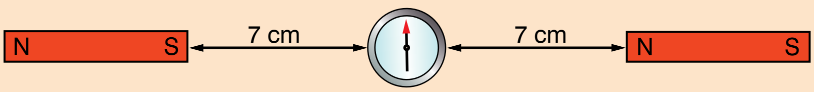


Figure 4.8 *How to set up the activity.*

Record the angle of deflection as you did in Activity 4.3.

Now remove the magnets and place a tin can over the compass. Put the magnets back in the same position as shown in Figure 4.9.

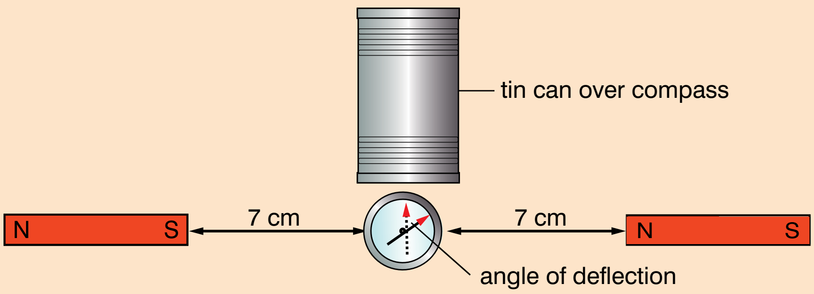


Figure 4.9 *Shielding the compass.*

Record the angle of deflection. Remove the magnets and place a second can over the first so that the compass is now shielded by two cans. Replace the magnets. Record the angle of deflection.

Repeat this procedure using first the plastic cups and then the paper cups. Place your results in a table like this.

Shielding material	Angle of deflection of compass
No shield	
One tin can	
Two tin cans	
One plastic cup	
Two plastic cups	
One paper cup	
Two paper cups	

Which material is the best magnetic shield?

Summary

- The Earth has a magnetic field that can be detected using a compass.
- Magnetic materials have atoms that act as tiny magnets which we call domains.
- When the domains are lined up, then the material is magnetic.
- If the domains are arranged randomly, then the material loses its magnetism.
- Some materials can be used as magnetic shields to protect equipment from becoming magnetised. Typical materials used are sheet metal, metal foam and plasma (ionised gas).

Review questions

1. Explain why a compass will show you which direction is magnetic north.
2. a) What is a magnetic domain?
   b) How can domains be used to explain what happens when a piece of steel becomes magnetised?
3. Describe how you could demonstrate magnetisation of iron filings or shredded steel wool.
4. What happens to the domains when a magnetic material is heated?
5. Describe how a nail can be magnetised.
6. Describe how you could find out which of a selection of materials was the best magnetic shield.

KEY WORDS

magnetic field *a region where a magnetic force may be exerted*

magnetic flux *a measure of the strength of a magnetic field over a given area*

4.2 Concepts of magnetic field

By the end of this section you should be able to:

- Describe a magnetic field.
- Perform and describe an experiment to demonstrate the existence of a magnetic field around a current-carrying wire.
- Sketch the resulting magnetic field lines pattern of a current-carrying wire.
- Apply the right-hand rule to determine the direction of magnetic field lines around a straight current-carrying wire.
- Calculate the magnetic field strength at a point due to a straight current-carrying wire.
- Sketch the magnetic field lines pattern of a current loop.
- Sketch the magnetic field lines pattern of a solenoid.
- Specify the polarity of a solenoid using the right-hand rule.
- Calculate the magnetic field strength at the centre of a solenoid.

What is a magnetic field?

A **magnetic field** is a region in which a magnetic force may be exerted. Put a compass down in a magnetic field and it will experience a force making it set in a particular direction.

A straightforward and strong magnetic field is the one in the gap of a horseshoe magnet. Any compass placed in the gap will feel a force turning it into the direction shown Figure 4.10a. The arrow represents the north end of the compass needle (that is, the end which normally points towards the north).

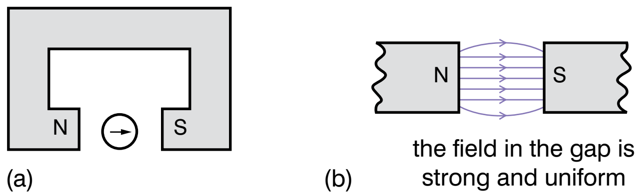


Figure 4.10 *The magnetic field in the gap of a horseshoe magnet.*

Figure 4.10b shows that same magnetic field as it is usually mapped in a way suggested by Michael Faraday (see page 143) in the 1820s. The lines are called lines of force or lines of **magnetic flux** (see page 142 for more about magnetic flux). They indicate the direction a small compass would set at any point, the arrow showing the way the north end of the needle points. Compare the field lines in Figure 4.10b with the behaviour of the compass in Figure 4.10a.

The field around a single bar magnet is as shown in Figure 4.11.

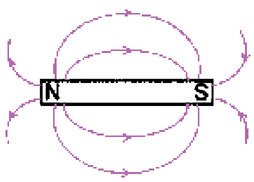


Figure 4.11 *The field around a single bar magnet.*

Notice how the lines of force start on the north pole of the magnet and end on the south, just as they did with the horseshoe magnet. Where the field is strongest (close to the poles), there the lines of force are closest together.

We sometimes use this idea and refer to the field strength as the 'magnetic flux density.' We shall learn more about this on page 142.

Activity 4.7: Magnetic fields in two dimensions

Place a strong bar magnet under a sheet of transparent material, as shown in Figure 4.12.

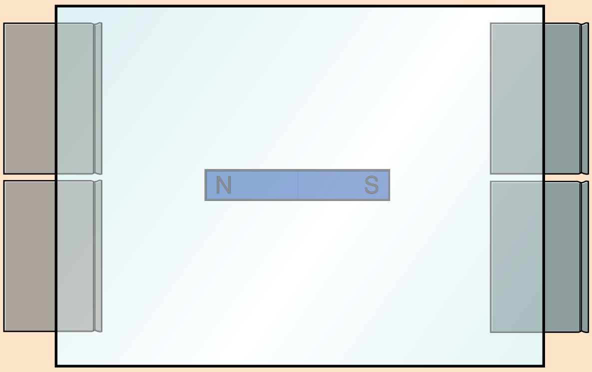


Figure 4.12 *Exploring magnetic fields in two dimensions.*

Make sure that the transparent sheet is level and then tap a beaker of iron filings so that they fall evenly over the surface. (Take care that you do not spill the filings as they will cause rust stains on clothing.) Draw the pattern you observe.

Now carefully put the iron filings back into the beaker. Remove the transparent sheet and place the bar magnet in the centre of a piece of paper. Draw round the outline of the magnet. Use a small compass placed near the magnet to plot the field lines around the magnet (look back to Figure 4.10a if you need to be reminded how to do this). Your plot should look like your drawing of the pattern you observed and Figure 4.11.

Activity 4.8: Magnetic fields in three dimensions

Place iron filings in the bottom of a glass jar and fill the remainder of the jar with salad oil. (As before, take care not to spill the iron filings.) If you do not have iron filings, then you can create some by rubbing two pieces of steel wool together rapidly.

Place a stopper in the top of the jar then shake the jar vigorously until the iron filings are distributed evenly throughout the container. Place the jar in the magnetic fields created by bar magnets outside the jar, with the south pole of one magnet opposite the north pole of the other, as shown in Figure 4.13.

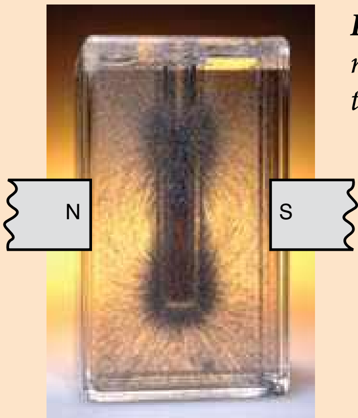


Figure 4.13 *Exploring magnetic fields in three dimensions.*

Allow time for the iron filings to align with the magnetic field, then draw a diagram of the field lines shown by the iron filings.

Magnetic field lines around a current-carrying wire

Magnetic fields are not only produced by metal magnets. An electric current will cause a compass needle to deflect. In other words, an electric current gives rise to a magnetic field. This magnetic field

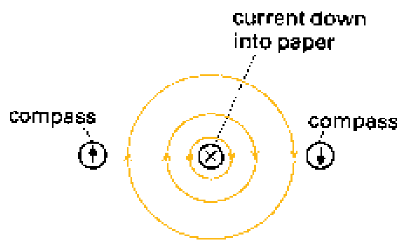


Figure 4.14 *The magnetic field around a straight current-carrying wire.*

can be represented as a series of circles centred round the wire, their spacing increasing with distance as the field gets weaker (see Figure 4.14).

This means that when the current is switched on, a small compass would tend to take up a position sideways onto the wire. You will show this in Activity 4.9.

Activity 4.9: Plotting the magnetic field lines around a current-carrying wire

Set up the apparatus as shown in Figure 4.15.

Switch on the current and then use a compass to plot the field lines in the same way as you did in Activity 4.7. Check that your results agree with Figure 4.14.

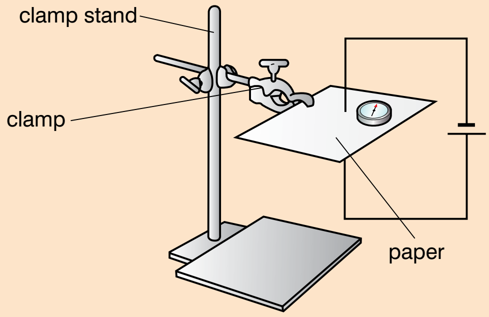


Figure 4.15 *How to set up the apparatus.*

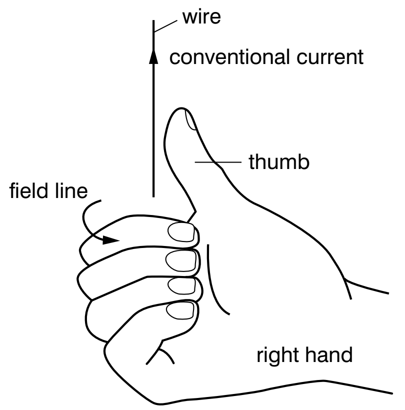


Figure 4.16 *Predicting which way the arrows will go using the right-hand rule.*

The right-hand rule

A way to predict which way the arrows go on the lines of force (that is, which way the north end of a compass will point) is shown in Figure 4.16.

Take your right hand (and not your 'left' one!) and hold the wire with your thumb going in the direction of the conventional current. The way your fingers then wrap round the wire is the way the field lines go.

Worked example 4.1

Use the right-hand rule to work out which way the field lines go in each of these diagrams.

The diagram shows which way the field lines go in each situation. Check that you can see why this is the case.

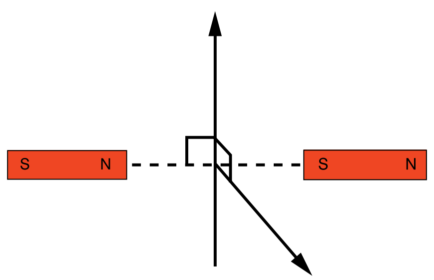


Figure 4.17 *The force on a current-carrying conductor.*

Magnetic field strength at a point due to a current carrying wire

We saw on page 127 that a current-carrying wire produces a magnetic field. We shall see later (on page 133) that when a current flows in a wire and there is a magnetic field perpendicular to the wire, as shown in Figure 4.17, the current-carrying wire will experience a force, F.

The strength of the magnetic field, which is given the symbol B, produced by a current-carrying wire, depends on:

- the force, F
- the current flowing through the wire, I
- the length of the wire, L.

The relationship is given by the formula

force = magnetic field strength × current flowing through the wire × length of wire

$F = B \times I \times L$

The units are newtons (N) for the force, amperes (A) for the current, metres (m) for the length and Teslas (T) for the magnetic field strength (also known as magnetic flux density).

Activity 4.10: Finding the magnetic field of a current loop

Set up the apparatus shown in Figure 4.18.

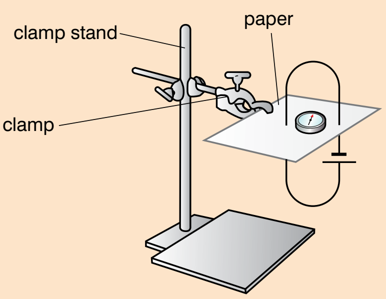


Figure 4.18 *Apparatus to find the magnetic field of a current loop.*

Plot the magnetic field lines using a compass as you did in Activity 4.9.

Magnetic field of a solenoid

A **solenoid** is a coil of wire that has a number of loops, as shown in Figure 4.19.

In Activity 4.10 you found that the magnetic field of a current loop is as shown in Figure 4.20 overleaf.

The currents in each side of the coil both contribute to the overall magnetic field. The field is strong in the centre of the coil but weaker outside the coil.

Worked example 4.2

Calculate the magnetic field strength when a current of 4 A flows in a wire 2 m long and produces a force of 8 N.

F (N)	*B* (T)	*I* (A)	*L* (m)
8	?	4	2

Rearrange $F = B \times I \times L$ so that B is on the left-hand side.

$B = \frac{F}{IL}$

Substitute the given values

$B = \frac{8}{4 \times 2} = 1\ \text{T}$

DID YOU KNOW?

The strength of the Earth's magnetic field changes all the time. When some bricks are made by heating in kilns, they become slightly magnetic and this magnetism stays in them as they cool down. If very sensitive equipment is used, this magnetism can be detected so that ancient buildings that are buried on an archaeological site can be found.

KEY WORDS

solenoid *a coil of wire in which a magnetic field is created by passing a current through it*

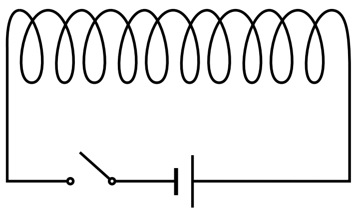


Figure 4.19 *A solenoid can be attached to a switch to allow current to be passed through.*

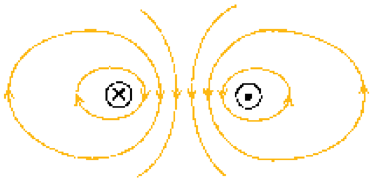


Figure 4.20 *The field created in a circular coil.*

In a solenoid, the magnetic field is made up from that produced by each turn. Inside the solenoid the field is strong and uniform, while outside it the field resembles that of a bar magnet as shown in Figure 4.21. ⊗ means current into paper, ⊙ means current out of paper.

Unlike a normal steel permanent magnet, the solenoid has a hole up the centre so any object which it attracts will tend to get pulled right into the middle. This is what happens in a washing machine, for example, at certain stages in its programme: a solenoid gets switched on, an iron rod is pulled into it against a spring, as shown in Figure 4.22, and the action turns a tap on to admit more water.

You can use the right-hand rule (see Figure 4.16) to work out which end of a solenoid is the north pole and which is the south. The thumb points to the north pole, if the fingers point in the direction of the current. Check that you can see why the solenoid shown in Figure 4.23 has its poles as indicated.

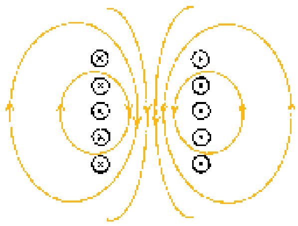


Figure 4.21 *The field produced by the current in a solenoid.*

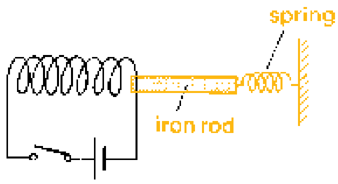


Figure 4.22 *When the solenoid is switched on by closing the switch, the iron rod is pulled into it.*

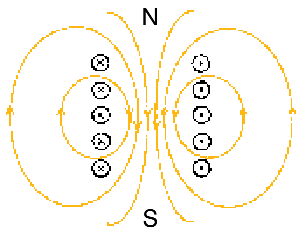


Figure 4.23 *Check that you understand why the poles on this solenoid are as indicated.*

DID YOU KNOW?

Permeability is the degree of magnetisation that a material obtains in response to an applied magnetic field. Magnetic permeability is typically represented by the Greek letter μ. The term was created in September 1885 by Oliver Heaviside.

In SI units, permeability is measured in the henry per metre (H/m). The constant value μ_0 is known as the magnetic constant or the permeability of free space, and has the exact (defined) value $\mu_0 = 4\pi \times 10^{-7}$ H/m.

Strength of magnetic field in a solenoid

The strength of the magnetic field in a solenoid (again given the symbol B) depends on:

- the number of turns of wire per metre of length, n.
- the permeability of free space, μ.
- the current flowing through the wire, I.

The formula is:

field strength = permeability of free space × number of turns per metre of length × current

$$B = \mu_0 n I \quad \text{where } n = \frac{N}{l}$$

Worked example 4.3

A solenoid has 1000 turns per metre and has a current of 2 A passing through it. Work out the field strength at its centre. The permeability of free space is $4\pi \times 10^{-7}$ H/m.

B (T)	n (m^{-1})	I (A)
?	1000	2

Substituting the given values we get

$B = 4\pi \times 10^{-7} \times 1000 \times 2 = 8\pi \times 10^{-4}$ T

Summary

- A magnetic field is a region in which a magnetic force may be exerted. If you put a compass down in a magnetic field, it will experience a force that makes it set in a particular direction.
- You can demonstrate the existence of a magnetic field around a current-carrying wire using the apparatus shown in Figure 4.24.

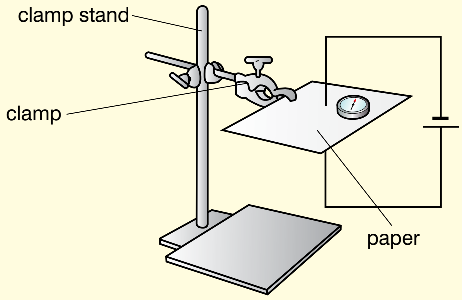


Figure 4.24

Switch on the current and then use a compass to plot the field lines.

- The magnetic field lines pattern of a current-carrying wire is as shown in Figure 4.25.

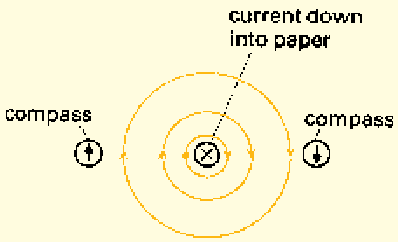


Figure 4.25

- You can apply the right-hand rule to determine the direction of magnetic field lines around a straight current-carrying wire. Take your right hand (and not your 'left' one!) and hold the wire with your thumb going in the direction of the conventional current. The way your fingers then wrap round the wire is the way the field lines go.
- The magnetic field strength at a point due to a straight current-carrying wire may be calculated using the formula:

force = magnetic field strength × current flowing through the wire × length of wire

$F = B \times I \times L$

- The magnetic field lines pattern of a current loop is as shown in Figure 4.26.

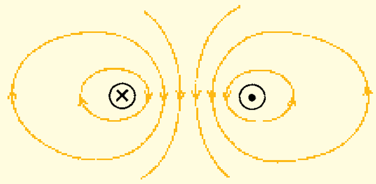


Figure 4.26

- The magnetic field lines pattern of a solenoid is as shown in Figure 4.27.

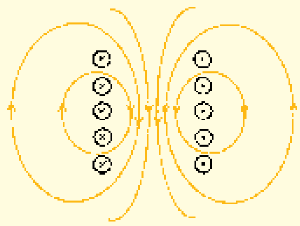


Figure 4.27

- You can work out the polarity of a solenoid using the right-hand rule and also remember that when you look at the solenoid from one end, a current flowing in a clockwise direction will behave like a south pole so a current flowing in an anti-clockwise direction will behave like a north pole (see Figure 4.28).

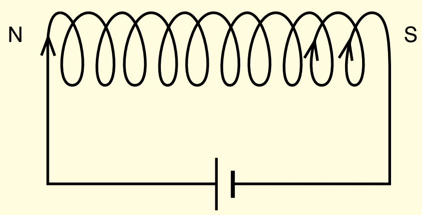


Figure 4.28

- You can calculate the magnetic field strength at the centre of a solenoid using the formula $B = \mu_0 nI$.

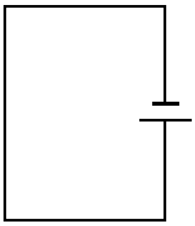


Figure 4.29

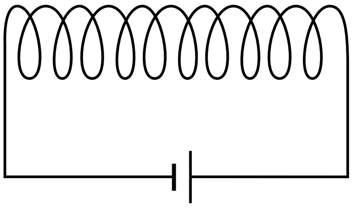


Figure 4.30

Review questions

1. a) What is a magnetic field?

   b) Describe an experiment that would demonstrate the existence of a magnetic field around a current-carrying wire.

2. Sketch the magnetic field patterns for:

   a) a straight current-carrying wire

   b) a current loop

   c) a solenoid.

3. Copy and complete the diagram in Figure 4.29 to show the direction of the field lines around the current-carrying wire.

4. Calculate the magnetic field strength when a current of 6 A flows in a wire 3 m long and produces a force of 36 N.

5. Copy the diagram in Figure 4.30. Use the right-hand rule to work out which pole is which in the solenoid.

6. Find the magnetic field strength at the centre of a solenoid with 5000 turns and current of 5 A. The permeability of free space is $4\pi \times 10^{-7}$ H/m.

4.3 Magnetic force

By the end of this section you should be able to:

- Describe the factors on which the force on a moving charge in a magnetic field depend.
- Demonstrate the relation $B = \frac{mv}{qR}$ from the fact that the centripetal force is provided by the magnetic force.
- Calculate the magnetic force acting on a moving charge in a uniform magnetic field.
- Determine the direction of a force acting on a moving charge using the left-hand rule.
- Demonstrate the existence of a force on a straight current-carrying conductor placed in a magnetic field.
- Derive the expression $F = BIl\sin\theta$ from $F = qvB\sin\theta$.
- Apply the left-hand rule to determine what will happen when current flows perpendicular to a uniform magnetic field.
- Calculate the magnitude and direction of force between two parallel current-carrying conductors in a uniform magnetic field.
- Define the SI unit ampere.

- Draw a diagram to show the forces acting on a rectangular current-carrying wire in a uniform magnetic field.
- Draw diagrams to show the action of a force on a simple d.c. motor and a moving coil galvanometer.

The magnetic force

We saw on page 129 that a current-carrying wire that has a magnetic field perpendicular to it will experience a force. We used the relationship:

force = magnetic field strength × current flowing through the wire × length of wire

$F = B \times I \times L$

So far all the moving charges have been part of a current that is flowing through a wire. There are occasions when individual charges travel through a magnetic field.

Even a single charge moving along constitutes an electric current. As it passes through a magnetic field at right angles to the field lines, you can apply Fleming's left-hand rule (see Figure 4.34 on page 135) to work out the direction of the force. There is just one point to watch out for – the current direction in Fleming's left-hand rule is that of conventional current, so if a negative electron is going towards the right that will be a conventional current to the left.

The 'IL' part of the $F = BIL$ expression for the magnitude of the force is replaced jointly by two factors: the size of the charge, q, and its velocity, v. Therefore, for a single charge, the force is given by

$F = Bqv$

Magnetic fields and the centripetal force

In Unit 1 you learnt that the centripetal force can be found using the equation:

$$F = \frac{mv^2}{R} \quad (1)$$

where F is the force, m is the mass of the particle, v is the velocity and R is the radius of the circle in which the particle is travelling.

A charged particle moving in a magnetic field will move in a circular path. We learnt in the last section that:

$F = Bqv$ (2)

If we make the two equations equal (that is, we assume that the force is the same in both), we can say that:

$$\frac{mv^2}{R} = Bqv$$

Activity 4.11: Deriving the relation $B = mv/qR$

With a partner, work out how you can rearrange

$$\frac{mv^2}{R} = Bqv$$

to give the relationship

$$B = \frac{mv}{qR}.$$

Worked example 4.4

A particle of mass 1.7×10^{-27} kg and velocity 8.0×10^6 m/s has a charge of 1 c and moves in a radius of 200 m. What is the magnetic field strength?

B **(T)**	***m*** **(kg)**	***v*** **(m/s)**	***q*** **(c)**	***R*** **(m)**
?	1.7×10^{-27}	8.0×10^6	1	200

Substituting these values

$$B = \frac{mv}{qR}$$

$$= \frac{1.7 \times 10^{-27} \times 8.0 \times 10^6}{1 \times 200}$$

$$= 6.8 \times 10^{-21}\ \text{T}$$

Activity 4.12: Demonstrating the motor effect

The motor effect can be demonstrated using the apparatus shown in Figure 4.31. A length of fairly stiff copper wire is bent into a square U-shape and is hooked over at the ends. It is allowed to dangle from the ends of the rest of the circuit as shown, so it can swing freely but the electric current can still be fed into and out of it.

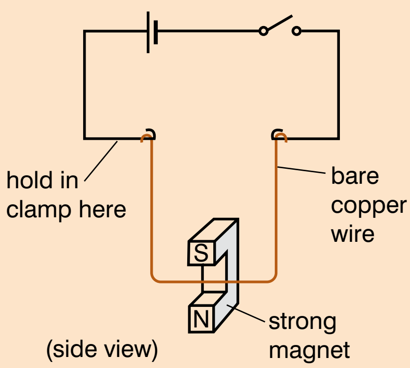


Figure 4.31 *The U-shaped copper wire can swing freely.*

As soon as you switch on the current, you will see a force acting in the direction shown on the wire carrying the current (see Figure 4.32). Reversing the current makes the wire move the opposite way, and so does reversing the direction of the magnetic field (by turning the magnet the other way up). Notice that the wire moves not in the direction of the magnetic field, but at right angles to it.

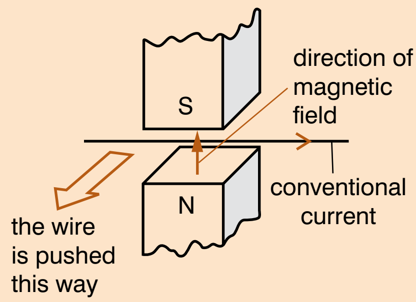


Figure 4.32 *The wire moves at right angles to the direction of the magnetic field.*

Explaining the motor effect

The arrows on magnetic lines of force show the direction of the force experienced by the north pole of a magnet. The south pole of the magnet will be pushed the opposite way.

Look at Figure 4.33a, which shows exactly the same situation as above except that only the magnetic field due to the current is shown.

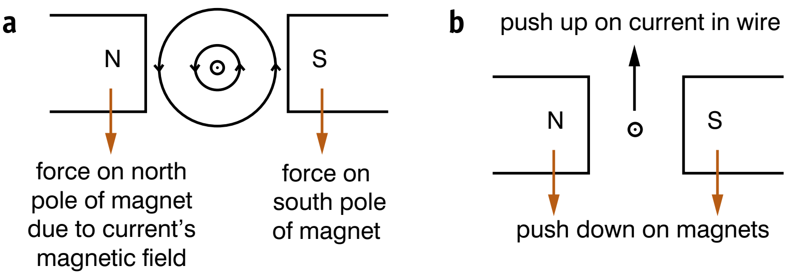


Figure 4.33 *A different way of explaining the motor effect.*

You will see that when the current flows, both magnets will be pushed downwards. (This force can be demonstrated by placing the magnets on a top pan balance.)

The magnets are usually fixed in position. By Newton's third law (which you learnt about in Unit 1), if they are being pushed down, the current in the wire will experience an equal sized push upwards (see Figure 4.33b). This is the motor effect.

Fleming's left-hand rule

To predict the direction of the movement produced by the motor effect, we can use Fleming's left-hand rule (see diagram). Hold the thumb and first two fingers of your left hand at right angles to each other. If the first **F**inger points along the magnetic field and the se**C**ond finger shows the **C**onventional **C**urrent, then the **TH**umb points in the direction of the **TH**rust (movement).

Factors that determine the force on a current-carrying conductor

The factors that determine the force are:

- the current
- the length of the wire
- the strength of the magnet.

We already know the relationship $F = BIL$ (see page 133). If the current is perpendicular to the field, the full BIL force is experienced. If the current is in the same direction as the lines of the magnetic field (**magnetic flux**), there will be no force.

In general, if there is an angle θ between the wire and the field (as shown in Figure 4.35), then:

$F = BIL\sin\theta$

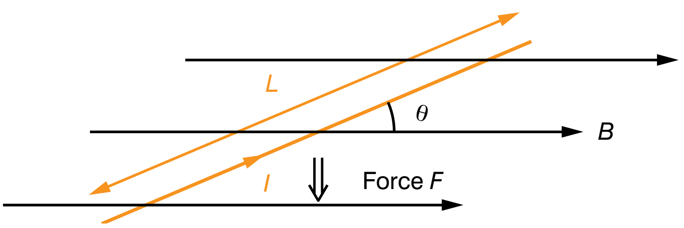


Figure 4.35 *Motor effect force.*

KEY WORDS

magnetic flux *a measure of the strength of a magnetic field over a given area*

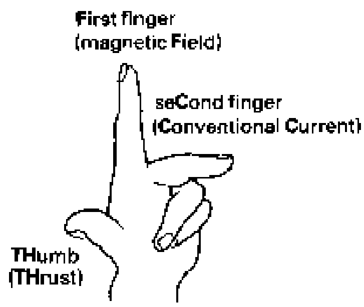


Figure 4.34 *Fleming's left-hand rule.*

DID YOU KNOW?

Sir John Ambrose Fleming (1849–1945) was a British electrical engineer and physicist. He invented the first thermionic valve or vacuum tube, the diode, in 1904 (see Unit 5). He also invented the left-hand rule. In 1932, he helped establish the Evolution Protest Movement. He did not have any children and, when he died, he left much of his money to charities, especially those that helped the poor. He was a good photographer and also painted watercolours and enjoyed climbing in the Alps.

Activity 4.13: Deriving $F = qvB\sin\theta$ from $F = BIL\sin\theta$

On page 133, we learnt that $F = Bqv$. You also know that $F = BIL\sin\theta$. Using these equations, derive the expression $F = qvB\sin\theta$.

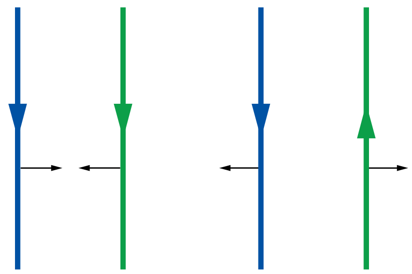


Figure 4.36 *Magnetic force between two parallel current-carrying conductors.*

Activity 4.14: Demonstrating the force between two parallel current-carrying conductors

Set up the two parallel current-carrying wires. First have the current in each wire flowing in the same direction and then reverse the direction of one of the currents.

Check that you understand why like currents attract.

Magnetic force between two parallel current–carrying conductors

Two parallel wires each carrying a current will interact with each other. If the currents are both flowing the same way, they attract one another; with currents going opposite ways they repel (see Figure 4.36).

The current in one wire creates a magnetic field that extends out to where the second wire is. The current in this second wire then experiences a force due to the motor effect.

Defining the ampere

This is the way that the size of a standard ampere has been fixed (you may well think of a current of 1 ampere as meaning that 1 coulomb of charge is flowing past every second, but in reality that is the way the coulomb is defined – an ampere is one of the fundamental units of the SI system). You will not need to remember the details, but if one ampere is flowing in each of two parallel wires 1 m apart in a vacuum, then the force on each wire due to the other will be exactly 2×10^{-7} N on every metre length. This rather strange figure was chosen because the ampere existed before the SI system was introduced, and this kept it the same size.

It means that in a standards laboratory, electric currents can be 'weighed' with a current balance. One of the wires is a circular coil held in a horizontal position. The other coil, just above it, takes the place of one of the pans on a pair of sensitive scales. When the current flows the same way in both coils, they attract. More weights have to be placed in the other pan of the scales to balance this – the current has been 'weighed'.

Force on a rectangular current-carrying wire

Consider a coil of length L (as in Figure 4.38) and N turns carrying a current I.

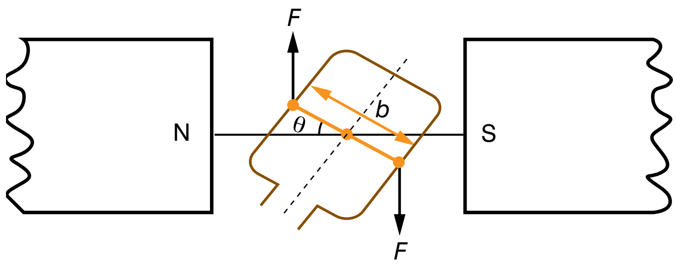


Figure 4.38 *A rectangular current-carrying coil.*

The plane of the coil makes an angle θ with the magnetic field.

If the current in the left-hand side of the coil is coming up out of the paper and that in the right-hand side is going down, the forces will be in the directions shown. The magnitude of each force F will be $BILN$.

These two forces provide a torque, a turning effect, on the coil.

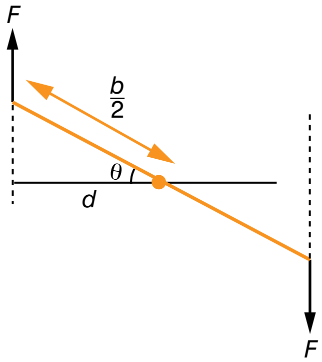


Figure 4.39 *Torque on a rectangular current-carrying coil.*

The total torque is the sum of the two moments. The distance the left-hand force acts from the pivot is d, which is $\frac{b}{2}\cos\theta$ (where b is the full width of the coil). Therefore its moment is $F \times \frac{b}{2}\cos\theta$.

The two forces combined give double this moment, which is $Fb\cos\theta$.

Now put in the size of force F ($= BILN$) and we get the torque to be $BILN\, b \cos\theta$.

Finally, note that $L \times b = A$, the area of the coil.

Therefore we end up with:

Torque on the coil = $BIAN \cos\theta$

When $\theta = 0$ the torque is a maximum. When $\theta = 90°$ the torque drops to zero – the two sides of the coil are still being pushed up and down, but the distance between those forces has fallen to 0.

The electric motor

Before proceeding further, check that you can convince yourself that the left-hand rule will predict the directions of the forces, as shown in Figure 4.40.

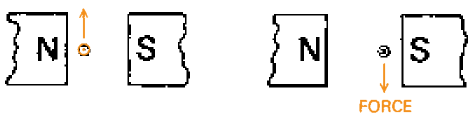


Figure 4.40 *Are the directions of the forces shown correctly?*

Figure 4.41 shows a coil carrying a current in a magnetic field.

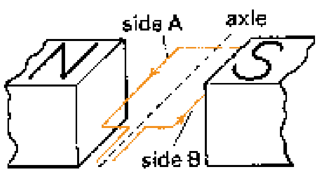


Figure 4.41 *The coil will do a quarter-turn then stop.*

DID YOU KNOW?

André-Marie Ampère (1775–1836) was a French physicist and mathematician who is generally regarded as one of the main discoverers of electromagnetism. The SI unit of measurement of electric current, the ampere, is named after him.

His father began to teach him Latin, until he discovered that Ampère preferred, and was good at, mathematical studies. Ampère, however, soon resumed his Latin lessons, so that he could read and understand the works of Euler and Bernoulli. Ampère later claimed that he knew as much about mathematics and science when he was eighteen as ever he knew; however, he also studied history, travel, poetry, philosophy and the natural sciences.

Figure 4.42 analyses the same coil by using Fleming's left-hand rule. It starts as in Figure 4.42a, the forces causing it to rotate. After a quarter of a turn 4.42b, the forces acting on the wires might distort the coil, but they will no longer turn it. If you pushed the coil round a bit more, the forces on the coil would simply return it to the upright position 4.42c.

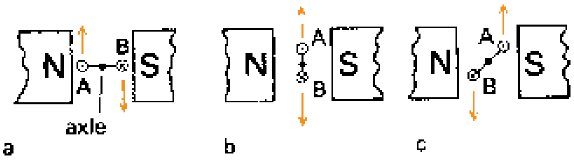


Figure 4.42 *Fleming's left-hand rule is used to predict the forces on the coil shown in Figure 4.41.*

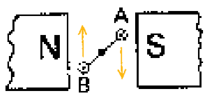


Figure 4.43 *The current through the coil has been reversed. Although the current in sides A and B of the coil itself will reverse its flow, the diagram always looks the same: on the l eft-hand side, and on the right.*

If by the time the coil reached the position of Figure 4.42c the battery leads to it could be reversed, so the current flowed the other way, the situation would become that of Figure 4.43 and the coil would continue to rotate.

To lead the current into the coil, and to reverse its direction automatically at the right moment, the coil ends up in two segments of metal called a split-ring commutator (Figure 4.44). The two wires from the battery end in brushes which press against each of the segments of the commutator.

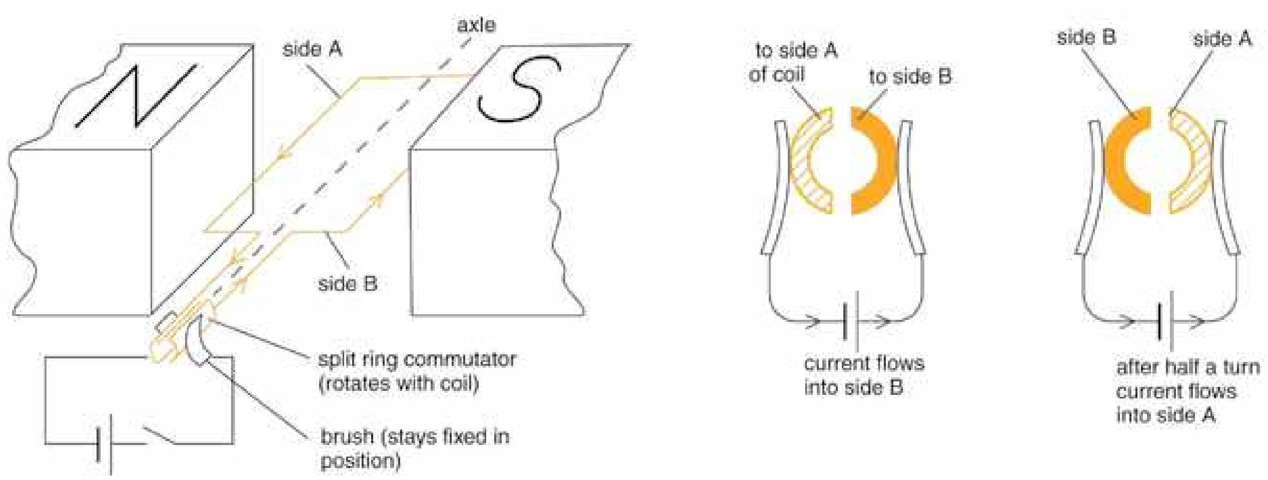


Figure 4.44 *How the direction of the current is reversed every half-turn of the coil.*

In cheap motors these brushes may be strips of springy metal, but in better ones they are usually blocks of carbon pressed against the commutator by springs. Sparking at the brushes tends to cause burns on the material used. With ordinary metals this is liable to lead to non-conducting corrosion, but carbon will oxidise to carbon dioxide gas, which will still leave a clean surface. Such brushes will need replacement from time to time as they wear down and burn away.

Moving coil galvanometer

The greater the current flowing around the coil of an electric motor, the more strongly it will try to turn. This suggests a way to measure the size of a current: let it flow through a motor, and make the coil try to turn while it is held back by a spring. The bigger the current, the further the coil will manage to stretch the spring.

This is the basis of the moving-coil galvanometer. The coil of the instrument is drawn in Figure 4.45a. The current can be fed into the coil and out again via the hairsprings at top and bottom; no commutator is needed because the rotation of the coil is restricted to just a fraction of a turn.

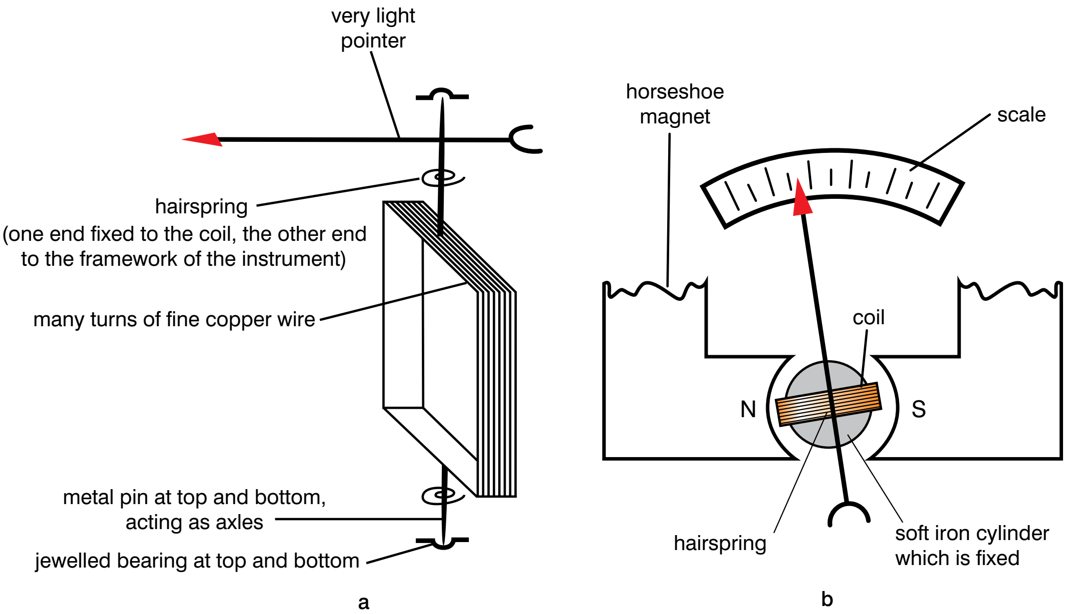


Figure 4.45 *Moving coil galvanometer.*

Figure 4.45b shows a view of the complete arrangement from above. The coil can rotate inside the gap of a steel horseshoe magnet whose pole pieces are curved. The soft iron cylinder which sits in the middle of the coil (but does not rotate with it) itself gets turned into a magnet because of the presence of the permanent magnet; one of its effects is to increase the strength of the field within the gap.

Its other effect is to give the instrument a linear scale. In the gap there is a radial field (think of how a small compass would set at that point), so as the coil rotates within the gap it always stays along the field lines. The 'cosθ' term does not appear in the torque, so the torque remains proportional to the current.

A galvanometer thus measures an electric current – 'galvanism' being an old name for current electricity. The greater the current round the coil, the more marked the motor effect is and the further the hairsprings are wound up.

A typical instrument is so sensitive that its pointer will be moved to the end of the scale by a current of perhaps 5×10^{-3} A; we say that it has a full-scale deflection of 5 mA. Even though copper is used for the windings of its coil, it consists of such a long length of so very thin wire that it may have a resistance as high as 50 ohms or more.

Summary

- The factors on which the force on a moving charge in a magnetic field depend are: the size of the magnetic field, B, the size of the charge, q, and its velocity, v. Therefore, for a single charge, the force is given by:

  $F = Bqv$
- The centripetal force is provided by the magnetic force and so you can equate the centripetal force equation and the magnetic force equations to give the relationship:

  $B = \frac{mv}{qR}$.
- You can determine the direction of a force acting on a moving charge using right-hand rule.
- You can demonstrate the existence of a force on a straight current-carrying conductor placed in a magnetic field, as shown in Activity 4.14.
- You can derive the expression $F = qvB\sin\theta$ from $F = BIl\sin\theta$, as shown in Activity 4.13.
- Apply the left-hand rule to determine what will happen when current flows perpendicular to a uniform magnetic field.
- The magnitude and direction of force between two parallel current-carrying conductors in a uniform magnetic field can be calculated when you know that the force on a wire in a magnetic field is $F = BIL\sin\theta$
- The SI unit ampere is defined as follows: if one ampere is flowing in each of two parallel wires 1 m apart in a vacuum, then the force on each wire due to the other will be exactly 2×10^{-7} N on every metre length.
- You can draw a diagram to show the forces acting on a rectangular current-carrying wire in a uniform magnetic field, as shown in Figure 4.45.

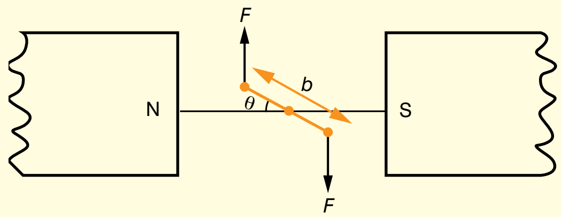


Figure 4.46

- You can draw diagrams to show the action of a force on a simple d.c. motor, as shown in Figure 4.46.

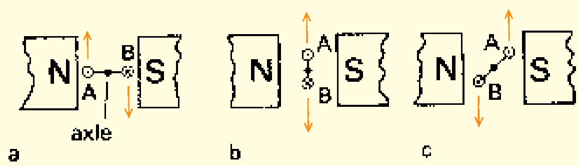


Figure 4.47

- You can draw diagrams to show the action of a force on a moving coil galvanometer in a similar way.

Review questions

1. Explain the factors on which the force on a moving charge in a magnetic field depends.
2. A particle of mass m carries a charge q and is travelling with a velocity v. It enters a region where there is a perpendicular magnetic field of flux density B.
   a) State the magnitude and direction of the motor effect force that will act on the particle.
   b) Explain fully why the path of the particle due to this force will be a circle.
   c) Show that the particle will be deflected into a circle of radius $r = \frac{mv}{Bq}$.

   d) Work out this radius for an electron in a vacuum entering a magnetic field of 0.02 T at a speed of 4.5×10^7 m/s. (The mass of an electron is 9.1×10^{-31} kg and it carries a charge of -1.6×10^{-19} C).

3. Explain the basic motor effect.
4. Explain Fleming's left-hand rule.
5. What are the factors that determine the size of a force on a current-carrying conductor?
6. Describe a demonstration of the force between two parallel current-carrying conductors.
7. Define an ampere.
8. a) What is the formula to find the force on a rectangular current-carrying coil?
   b) What is the force on a coil in a magnetic field of 0.2 T, with a current of 1 A, an area of 0.025 m^2 and 100 turns?
9. Describe how a basic electric motor works.
10. Describe how a moving coil galvanometer works.

4.4 Electromagnetic induction

By the end of this section you should be able to:

- Define magnetic flux and its SI unit.
- State Faraday's law of induction.
- Perform simple experiments that demonstrate an induced e.m.f. caused by changing magnetic flux.
- State Lenz's law.
- Indicate the direction of induced currents, given the direction of motion of the conductor and the direction of a magnetic field.
- Describe the factors that affect the magnitude of induced e.m.f. in a conductor.
- Describe the link between electricity and magnetism.
- Apply Faraday's law to calculate the magnitude of induced e.m.f.
- Define inductance and its SI unit.
- Distinguish between self- and mutual inductance.
- Apply the definition of inductance to solve simple numerical problems.
- Explain the action of the simple a.c. generator.
- Compare the actions of d.c. and a.c. generators.

Worked example 4.5

Find the magnetic flux when a magnetic field of strength 2 T covers an area of 2 m^2.

Ø (Tm^2)	**B (T)**	**A (m^2)**
?	2	2

Substituting the values we get:

magnetic flux = 2 T × 2 m^2
= 4 T m^2

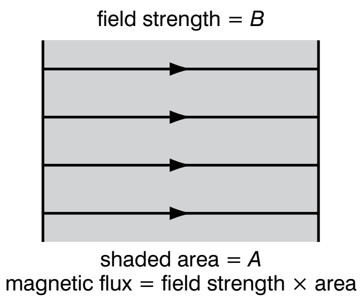


Figure 4.48 *Magnetic flux.*

KEY WORDS

electromagnetic induction *the production of voltage across a conductor moving through a stationary magnetic field*

induced e.m.f. *voltage produced by electromagnetic induction*

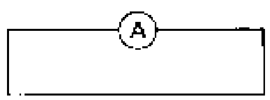


Figure 4.49 *A strange circuit.*

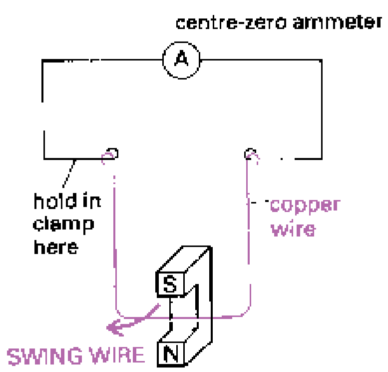


Figure 4.50 *How to demonstrate the dynamo effect.*

- Draw a diagram of a transformer.
- Give a simple explanation of the principles on which a transformer operates.
- Identify that, for an ideal transformer, $P_{out} = P_{in}$.
- Show that, for an ideal transformer, $V_s/V_p = N_s/N_p$.
- Apply the transformer formulae to solve simple problems.

Magnetic flux and magnetic field strength

We have already used the symbol *B* for magnetic field strength. Magnetic field strength is also magnetic flux density, the magnetic flux per unit area. So:

magnetic flux = magnetic flux density (magnetic field strength) × area

In symbols, we write this as:

$\varnothing = B \times A$

The unit for magnetic flux is T m^2 because *B* has the unit T and *A* has the unit m^2.

Electromagnetic induction

Figure 4.49 shows a strange circuit. It is just a loop of wire. To see if any current is flowing round it, a sensitive ammeter has been inserted. Let us suppose it is a centre-zero instrument, that is one in which the undeflected pointer is positioned in the middle of the scale rather than at the left-hand end of it, so whichever way a current happens to flow we can detect it.

An understandable first reaction is that we are wasting our time. There is no battery in the circuit to make the charge flow, no chemical energy being released to create the electrical energy.

Electromagnetic induction, also known as the dynamo effect, is one way to cause a current to flow round that circuit. Figure 4.50 shows a way to demonstrate it. We have that apparently pointless circuit with which we began.

Laws of electromagnetic induction

Faraday's law

The dynamo effect means that a voltage will appear whenever a conductor cuts through a magnetic field. To see what this means look at Figure 4.51, which shows an end-on view of a wire being moved between the poles of a magnet.

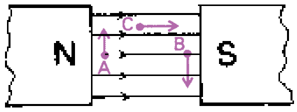


Figure 4.51 *An end-on view of a wire being moved between the poles of a magnet.*

At A the wire is cutting through the field and so a voltage is created in it. The same thing happens at B, but because the wire is cutting through the field in the opposite direction the voltage will be the other way round. At C, however, even though the wire is moving in the magnetic field no lines of force are being cut, and so no voltage is generated.

Careful observations reveal the following points about the dynamo effect:

1. It occurs when a conductor cuts through magnetic flux lines. What is produced is an e.m.f. in volts, so the wire behaves as if there was a battery in it. The actual current generated then depends on whether the circuit to which it is connected has a high resistance or a low one:

   $(I = \frac{V}{R})$.
2. **The size of the e.m.f. in volts is proportional to the rate at which the conductor is cutting through flux lines**. This is known as Faraday's law. Thus to generate double the e.m.f. the conductor must cut through twice as many flux lines each second. One way to do this is to move the conductor through the field at twice the speed. An alternative way is to keep the speed of the conductor the same but have the lines of flux twice as close together; in other words, double the strength of the magnetic field.

The dynamo effect is also called electromagnetic induction, and you may sometimes hear the voltage referred to as the **induced e.m.f**.

There are two laws of e.m.f. These are:

1 Faraday's law (about magnitude of the induced current).

2 Lenz's law (about direction of the induced current).

These two laws can be summarisedby the equation

$\varepsilon = \frac{-\Delta Ø}{\Delta t}$ where ε is induced e.m.f.
$\Delta Ø$ is charge on magnetic flux
Δt is change in time

Activity 4.15: Demonstration of the dynamo effect

When the wire is in the magnetic field but is not moving, nothing happens. Set up the apparatus as shown in Figure 4.50, take hold of the copper wire and swing it through the field of the magnet. While you are moving it through the field, the pointer of the ammeter will deflect to one side showing that a current is being generated. Then try moving the wire in the other direction through the magnetic field, and, for an instant, the ammeter will deflect the opposite way.

DID YOU KNOW?

Michael Faraday (1791–1867), a British chemist and physicist, contributed to the fields of electromagnetism and electrochemistry. He received little formal education yet he was one of the most influential scientists in history. Some refer to him as the best experimentalist in the history of science.

He established the basis for the electromagnetic field concept and made many contributions to chemistry, including inventing an early form of the Bunsen burner.

Activity 4.16: Faraday's law

Spin a weighted bicycle wheel. Bring a magnet close to the rim, which must be non-ferrous. What happens?

Now spin a metal disk (make it out of a non-magnetic material such as aluminium roofing metal). Bring a strong magnet near the disk. What happens?

Discuss the observations and try and explain them using Faraday's law.

Lenz's law

By swinging the wire of Figure 4.50 through the magnetic field you can generate a current which, in principle at least, could pass through a heating coil to boil some water. Electrical energy has then been converted to thermal energy, but where did the electrical energy appear from in the first place? The principle of conservation of energy demands that it has not just come from nowhere, but what has lost energy so the charge can gain it?

To answer this important question, consider this. A wire has been wound into a solenoid, and is now the magnetic flux lines that are cutting through the conductor because it is the magnet that is moving rather than the wire. The effect, however, is the same: an e.m.f. is induced.

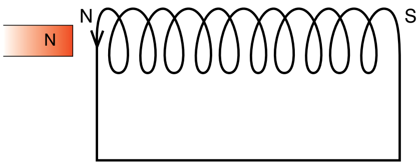
We know there are three ways to double the voltage that is generated.

1 Move the magnet in at twice the speed (Faraday's law).

2 Use a magnet that is twice as strong (Faraday's law again).

3 Have twice as many turns on the coil. Each turn behaves as if there was a little battery induced in it, so in effect you now have twice as many batteries in series in the circuit.

None of this gives any indication as to the source of that energy, however.



As the magnet is being moved towards the solenoid, the e.m.f. generated by the dynamo effect causes a current to flow round the circuit. During this time the solenoid too will behave as a magnet, and there is the clue as to where the energy is coming from.

There is a general principle by which the direction of the current produced by the dynamo effect may be predicted. It is known as Lenz's law:

the direction of the induced current is such as to oppose the change that is causing it.

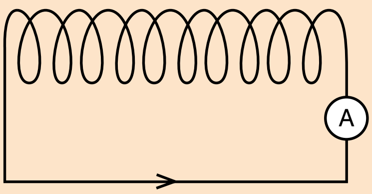
Lenz's law applies to any situation in which the dynamo effect occurs. To illustrate what it means, we will apply it to our magnet being inserted into the solenoid.

The north pole of the magnet is coming towards the solenoid, and this is the 'change' that is causing the dynamo effect here. Lenz's Law is saying that the direction of the current generated is sure to make the top of the solenoid a magnetic north pole too, so by repulsion it opposes the approach of the magnet.

The question was: where does the energy come from? All the time you are bringing the magnet up to the solenoid there is a force of repulsion. You must do work as you push the magnet against this opposing force. To generate the current you have got to release chemical energy by 'burning' some food as you push the magnet; and that is where the energy has come from.

Activity 4.17: Illustrating Lenz's law

Set up a solenoid as shown.



Bring the poles of a magnet towards each end of the solenoid. What happens to the needle on the ammeter?

When you remove the magnet again, the current flows in the other direction. The south pole now at the top of the solenoid will attract the north pole of the magnet you are trying to remove, and so you must pull the magnet out against this opposing force. Once more you must do work, and once more the drop in your reserves of chemical energy is matched by the appearance of electrical energy.

Activity 4.18: Lenz's law in action

Drop a strong small magnet through a thick copper pipe that is about 1 m long. Drop a rock through another similar tube. What happens?

Now try putting about 1000 turns of wire around the copper pipe and attach them to a light bulb as shown in Figure 4.52. Drop the magnet through the pipe again. What happens?

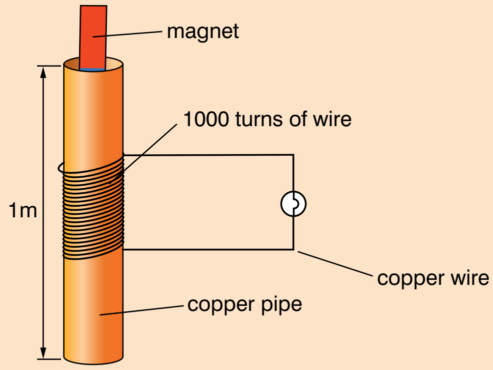


Figure 4.52 *Apparatus set up.*

Explain your observations using Lenz's law.

Activity 4.19: Magnets producing movement

Suspend a non-ferrous ring from two points. Let it swing freely and stop it. Put a magnet on a stick and push it through the ring. What happens?

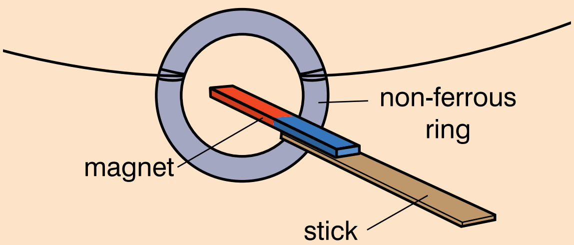


Figure 4.53 *Apparatus set up.*

Inductors

An inductor is essentially an electromagnet. The iron core that becomes magnetised when you send a current through a solenoid is not part of the actual circuit, but instead relies on the magnetic field associated with the electric current flowing round the coil.

Activity 4.20: Factors that affect the magnitude of an induced current in a conductor

In a small group, design a poster to summarise what you have learnt about the factors that affect the magnitude of an induced current in a conductor.

Think of some practical examples of the dynamo effect you use in everyday life.

Activity 4.21: The relationship between the motor effect and the dynamo effect

Discuss how the motor effect and the dynamo effect are related. If you can use the left-hand rule to work out the direction of motion in the motor effect, which rule, that you have already met, could you use to work out the direction of induced current in the dynamo effect?

At first sight the iron core makes no contribution whatsoever to the behaviour of the circuit itself, and the current is determined solely by the resistance of the coil and the voltage of the supply. This is true when a constant smooth d.c. is flowing. Whenever that current *changes*, however, the magnet makes its presence felt – when you first switch the current on, when you switch the circuit off, or all the time if the current is an alternating one.

The dynamo effect comes into play whenever the magnetic flux through it is changing.

a Inductor **b** Inductor with iron core

Figure 4.54 *a Air-cored inductor b iron-cored inductor*

The symbol for an air-cored inductor is shown in Figure 4.54a. The variation drawn in Figure 4.54b represents an inductor with an iron core, in which case the effect is considerably enhanced. This is the first circuit symbol you will have met which shows something which is not a direct part of the electrical circuit.

All electromagnets possess **inductance**. If the device is intended to act as a circuit component rather than to pick up pins, however, it is more efficient to wind the coil round a closed iron core which acts as a ring magnet.

Inductance is defined as the property in an electrical circuit where a change in the electric current through that circuit induces an e.m.f. that opposes the change in current.

KEY WORDS

inductance *the property in an electrical circuit where a change in the electric current through that circuit induces an e.m.f. that opposes the change in current*

self-inductance *the ratio of the electromotive force produced in a circuit by self-induction to the rate of change of current producing it*

Self-inductance

When you first turn on, the current starts to build up, which causes the magnetic flux through the coil to increase. This induces a voltage in the coil which, by Lenz's law, is in a direction such as to oppose what is causing it – in other words, it is a back e.m.f. which acts against the battery in the circuit to delay the build up of current. It is as if the inductor has provided the current with a kind of inertia.

The magnitude of the effect is specified by the coil's **self-inductance**, *L*. The induced e.m.f. depends on the rate at which the current in the coil is changing. We say the:

induced e.m.f. = a constant × the rate of change of the current.

induced e.m.f. = *L* × the rate of change of current through the coil

In sumbols, this is written $\varepsilon_{\text{ind}} = \dfrac{L\Delta I}{\Delta t}$

The value of that constant depends on the coil itself – its geometry, the number of turns on it, what its core is made from – and represents its self-inductance. A large value of *L* means that, for a given rate of change of current through it, a big e.m.f. will appear in it.

Since the rate of change of current will be measured in A/s, the SI unit of inductance will be V/A s, which is named the henry, H.

When a circuit comprising a battery of e.m.f. *V*, an inductor (*L*) and a resistor (*R*) in series is turned on, Figure 4.55 shows how the current builds up to its final steady value of $\frac{V}{R}$.

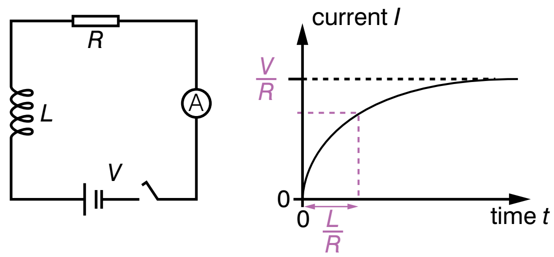


Figure 4.55 *Circuit comprising a battery, an inductor and a resistor in series, and graph showing self-induced e.m.f.*

The current is turned on at time $t = 0$. It then rises to its final value of $\frac{V}{R}$ in the form of an inverted exponential curve, reminiscent of the build-up of charge as a capacitor fills up through a resistor (see Unit 2).

With a capacitor, the time constant of the circuit was given by CR and represented the time taken for the curve to rise to about 63% of its final value. Here the time constant is given by $\frac{L}{R}$.

When you switch such a circuit off, the current cannot then die down exponentially to zero because the circuit has been broken. The current has to stop flowing almost instantaneously. That induces a momentary huge e.m.f. – this time a forward one, trying to keep the current flowing. If the inductor had given the current a kind of inertia as it built up, this is now like a hammer blow just for an instant. It can easily cause a spark between the contacts of the switch as they separate.

Mutual inductance

It is worth mentioning at this stage that two coils in separate circuits can show what is called not self-inductance but **mutual inductance**. When the current is changing in one circuit, its changing magnetic field cuts through the other circuit and induces a voltage in it.

The faster the current in one changes, the greater the e.m.f. induced in the other. This linkage between the two circuits is described as mutual inductance. It will be small if the two circuits are well spaced apart; it will be large if a coil in one circuit is wound round a coil in the other circuit.

The mutual inductance M of a pair of circuits is defined by:

e.m.f. induced in one circuit = M × (rate of change of current in the other circuit)

In sumbols, this is written $\varepsilon_{mut} = \frac{M\Delta I_2}{\Delta t}$

The unit of M will also be the henry, H.

A simple a.c. generator

It is difficult to generate a continuous current by swinging a wire in a magnetic field. Whichever way you move the wire you soon come out of the magnetic field, which means the wire is no longer cutting through flux lines.

KEY WORDS

mutual inductance *the ratio of the electromotive force in a circuit to the corresponding change of current in a neighbouring circuit*

Worked example 4.6

Work out the induced e.m.f. in a circuit linked to a second circuit of inductance 3.0×10^{-6} H when the rate of change of current in the second circuit is 3 A/s.

ε_{mut} (V)	M (H)	$\Delta I_2/\Delta t$ (A/s)
?	3×10^{-6}	3

Substitute the given values

induced e.m.f. = 3.0×10^{-6} H $\times$ 3 A/s = 9×10^{-6} V

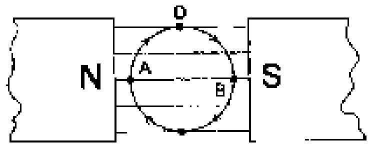


Figure 4.56 *An end-on view of a wire being moved in a circle in a magnetic field.*

Figure 4.56 shows an end-on view of such a wire.

The wire is going round at a steady speed, so it is cutting the lines of force at a maximum rate at points A and B. At these moments the voltage induced in the wire is at its greatest, one way round at A and the other way round at B.

At stages C and D in the rotation, although the wire is moving, no lines of force are being cut, so at those instants the voltage is zero. The graph (Figure 4.57) shows what will happen in the course of one complete revolution, starting at C.

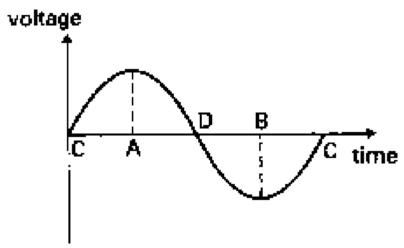


Figure 4.57 *The voltage induced in the wire in Figure 4.56 during one complete revolution.*

Suppose you have not a single wire but a rectangular coil which may be rotated about an axle. As the left-hand side moves up through the magnetic field, the right-hand side will move down. The voltages induced at that moment have been represented in Figure 4.58 by cells. Notice how, although they are opposite ways round, the two combine to pump current round the coil rather than cancelling each other.

A practical arrangement by which the coil may be rotated and yet joined to a circuit is shown in Figure 4.59. It is very similar to the simple electric motor discussed earlier, but this time the coil has to be turned rather than turning by the motor effect.

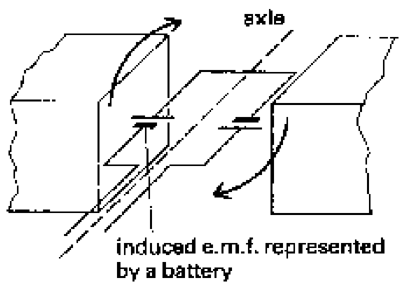


Figure 4.58 *The induced voltages are represented by batteries. The two combine rather than cancelling each other out.*

The connection to the outside circuit is again made by sliding brush contacts. It is not necessary to reverse those connections every half a turn, however, so the commutator is replaced by a pair of slip rings. Notice carefully how the two slip rings are joined to the coil. They are metal rings whose centre is the axle, and they spin round with the coil.

The output of this dynamo will be in the form of an alternating current (the continuous line in Figure 4.60) whose frequency is determined by the rate at which the coil is being rotated. Suppose the coil takes $\frac{1}{20}$ s to complete one revolution. It will then turn 20 times every second, so the frequency of the alternating current (a.c.) will be 20 Hz.

If the dynamo is turned at twice the rate, two things will change: 1) the coil turns in half the time, so the frequency of the a.c. will double; 2) by Faraday's law, since it cuts through the magnetic field twice as fast, the induced voltage will be doubled and so will the current. The dotted line on Figure 4.60 shows this.

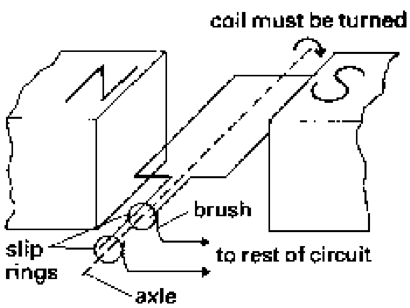


Figure 4.59 *How the coil is joined to the circuit.*

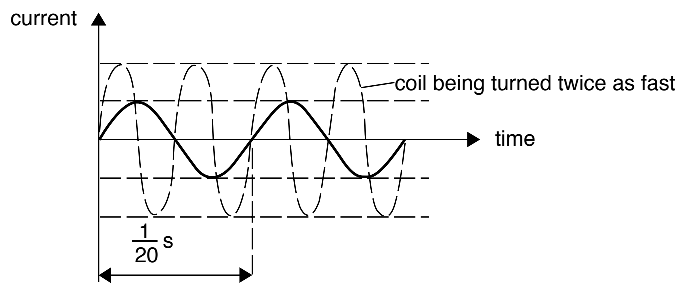


Figure 4.60 *The output of the dynamo in Figure 4.59.*

How transformers work

Look at Figure 4.61. The top half shows an alternating current supply feeding a solenoid. This causes the iron rod through the solenoid to be magnetised first one way then the other.

In effect the lower coil is having a magnet forever plunged into it one way round then the other. By the dynamo effect this will produce a voltage in the coil, and the current that results can be sufficient to light the lamp.

Electrically the two circuits are quite separate. It is the iron rod acting as a magnet that links the two circuits: the first circuit magnetises it, and by the dynamo effect the second circuit detects the arrival of this magnet. This how a **transformer** works.

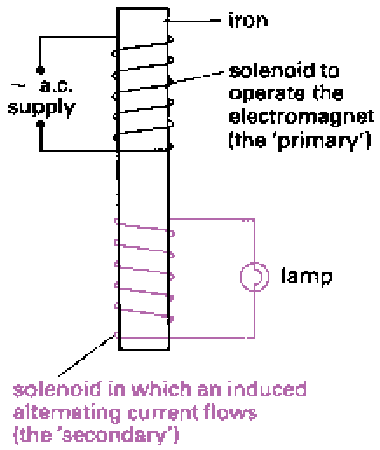


Figure 4.61 *The principle of the transformer.*

The iron is usually in the shape of a closed ring. This gives the arrangement drawn in Figure 4.62. Note the names: the input winding, intended to magnetise the iron core, is called the primary; the output winding, into which an e.m.f. is induced by the dynamo effect, is the secondary.

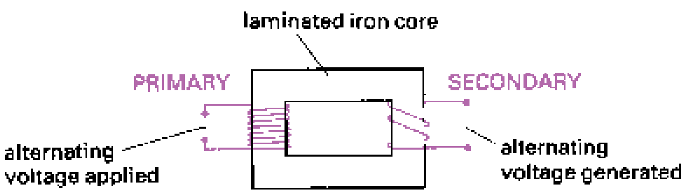


Figure 4.62 *The usual arrangement in a transformer.*

A word of explanation is needed about the laminated iron core. The word 'laminated' means 'made out of flat sheets'. The iron core is not a solid chunk of iron: it is made from a large number of separate pieces, each only a millimetre or so thick, as Figure 4.63 shows.

Iron is used for the core because of its magnetic properties, but it is also a metal and metals conduct electricity. Therefore there is one more electrical circuit around: the iron of the core.

With all those magnetic lines of force building up then collapsing back again, eddy currents are induced in the core itself, a bit like water swirling and splashing round in a bowl. These eddy currents

Activity 4.22: Build a simple a.c. generator

Use the information on page 148 to build a simple a.c. generator.

Activity 4.23: Compare the actions of a.c. and d.c. generators

In a small group, draw up a table that compares the actions of a.c. and d.c. generators.

KEY WORDS

alternating current (a.c.) *an electric current that periodically reverses direction*

transformer **a** *device that transfers electrical energy from one circuit to another, usually with a change of voltage*

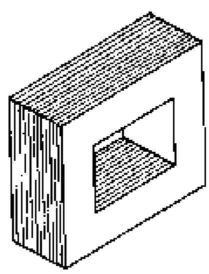


Figure 4.63 *Each of the separate pieces of laminated core is only about 1 mm thick.*

Eddy currents

We usually picture an electric current as resembling water flowing along a pipe. There are other forms of water currents around, however, such as those you will generate if you swirl water around in your bath.

The same applies to electric currents too. If a metal plate or a three-dimensional metal object is located where magnetic fields are changing rapidly, such *eddy currents* are likely to be induced.

Sometimes this represents a waste of energy and unwanted heat being produced. In the iron core of a transformer, the problem is kept to a minimum by laminations.

Worked example 4.7

The 110 V mains is applied to the primary of a transformer consisting of 800 turns. The secondary has 200 turns. What is the output voltage?

The output voltage will be less than 110 V, because the secondary has the smaller number of turns. The turns ratio is 800:200, which cancels down to 4:1. The output voltage will therefore be a quarter of the input voltage; that is, 110 divided by 4 which comes to about 28 V.

will generate heat in the iron, and this is not welcome on two counts: it represents energy being fed in at the primary which does not emerge at the secondary, and the heat produced may cause overheating.

To stop water from circulating in a bowl, a series of baffle plates could be inserted to block the flow paths of the water. In the transformer core the electrical equivalent is done. Each of the flat iron plates is insulated from its neighbours by paint, varnish or sometimes just rust (Figure 4.64).

Figure 4.64 *Why the core is laminated.*

Step-down and step-up transformers

The alternating voltage applied to the input drives a current round the primary to magnetise the core. As the state of magnetisation of the core changes, an alternating voltage will be induced in the secondary. There is no reason why these two voltages should be the same size and, in fact, the main purpose of transformers is to change the size of a voltage.

The one shown in Figure 4.65 is a **step-down transformer**: it steps the voltage down so that low voltage equipment can be run from the mains.

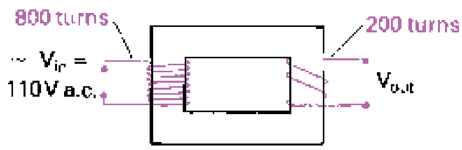


Figure 4.65 *A step-down transformer.*

You can predict the output voltage of a transformer (V_{out}) like this:

$$\frac{V_{out}}{V_{in}} = \frac{\textbf{number of turns on the secondary } (N_s)}{\textbf{number of turns on the primary } (N_p)}$$

In other words, the turns ratio is the same as the ratio of the two voltages. The lower voltage is associated with the coil that has the smaller number of turns. Remember that both the input voltage and the output voltage will be alternating ones.

A **step-up transformer** works the other way. It is the secondary that has the greater number of turns, so the 110 V of the mains might be stepped up to a few kilovolts, for instance. This happens inside a television set, and explains why you must never switch it on if the back is removed. It is far more dangerous than the mains.

Figure 4.66 shows the circuit symbol for a transformer. Assuming that the connections on the left are to the primary coil, (a) is a step-down transformer and (b) is a step-up one. It is normal to indicate which coil has the fewer turns, though no attempt is made to suggest the actual turns ratio.

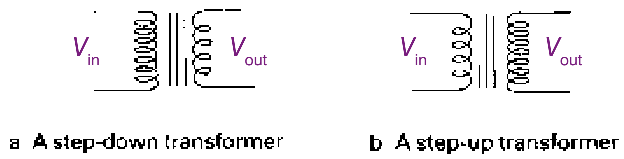


Figure 4.66 *The circuit symbols for (a) step-down and (b) step-up transformers.*

Activity 4.25: Lenz and Faraday and decaying fields

Discuss the following question in small groups.

The primary coil of a transformer is connected to a battery, a resistor and a switch. The secondary coil is connected to an ammeter. When the switch is closed, does the ammeter show:

a) zero current

b) a non-zero current for an instant

c) a steady current?

Some uses for such currents are:

- spark plugs in a car
- heart starting paddles in a hospital
- electric fences for animals.

Research one or more of these applications.

The ideal transformer equation

There are two ways in which a transformer may heat up:

1. Eddy currents may be induced in the iron core. Laminating the core has very nearly solved this problem.
2. The resistance (in ohms) of the windings of the coils themselves could cause heating. However, the wires are made from a low-resistance metal, such as copper, and are thick enough to cope with the currents expected.

Thus, with a well-designed transformer being used for the job for which it is intended, its efficiency is very close indeed to 100%. This means that almost all the energy fed into the primary will emerge at the secondary. An ideal transformer has an energy transfer of 100%.

Consider a transformer run off the mains, which, to keep the numbers simple, we will imagine to be not 110 V but 100 V. The transformer is a step-down one with a turns ratio of 10:1 (so the number of turns on the secondary is only a tenth that of the primary). The voltage at the output, therefore, will be 10 V.

Activity 4.24: A flyback transformer

Use a flyback transformer from an old TV. Connect a small bulb in series with a 12 V source and a flyback transformer. What happens? Why?

Activity 4.26: Transient electric current

Make a high number of windings from a broken transformer's wires. Attach a small bulb to the coils. Let a magnet fall through the coils. What happens to the bulb as the magnet falls through the coils?

KEY WORDS

step-down transformer *one in which the output a.c. voltage is less than the input*

step-up transformer *one in which the output a.c. voltage is greater than the input*

Suppose the output is fed to a circuit of resistance 5 Ω. The 10 V will supply an alternating current of 2 A to the circuit (since $I = \frac{V}{R}$). How large will the current 'I' be that is drawn from the mains (Figure 4.67)?

You may wish to look back to Unit 3 at this point!

***Figure 4.67** How large is the current being drawn from the mains?*

To answer this, use conservation of energy and recall that the power $P = VI$. The power at the output of the transformer is 10 V × 2 A, which is 20 W. In other words, 20 J of energy every second are being taken from the output of the transformer.

Assuming 100% efficiency (an ideal transformer), exactly 20 J of energy must be supplied by the mains in that time. Again using $P = VI$, we get:

$20 \text{ W} = 100 \text{ V} \times I$

$\therefore\ I = 0.2 \text{ A}$

In general, therefore, for a transformer:

$$V_{out} \times I_{out} = V_{in} \times I_{in}$$

You may sometimes see this same relationship expressed in the form:

$$\frac{I_{in}}{I_{out}} = \frac{V_{out}}{V_{in}} = \frac{N_s}{N_p}$$

where V_{out} is the alternating voltage produced in the secondary coil, and V_{in} is the alternating voltage which is applied to the primary coil, for N_s and N_p.

Activity 4.27: Show that for an ideal transformer $P_{out} = P_{in}$

Use the information here to show that, for an ideal transformer, $P_{out} = P_{in}$.

Worked example 4.8

What would the output voltage be if the secondary coil had 250 turns and the primary coil had 1000 turns, and the input voltage were 100 V?

N_s	N_p	V_{in} (V)	V_{out} (V)
250	1000	100	?

Rearrange the formula to give $V_{out} = \frac{N_s \times V_{in}}{N_p}$

Substitute given values: $V_{out} = 250 \times \frac{100}{1000} = 25$ V

Activity 4.28: Build an a.c. motor

Build either a toothpick motor or the cork motor. Your teacher will give you instructions about how to do this.

Measure the rotation rate. You should aim for as fast a rotation rate as possible as the fastest rotators convert the electrical energy to kinetic energy most efficiently.

Your motor must have a small piece of reflecting material that bounces back a flashlight beam. The beam should hit a simple photo transistor circuit connected to a buzzer. Count the number of buzzes in 2 minutes.

You may need to look at Unit 5 for some help with this.

Summary

- Magnetic flux = magnetic flux density (magnetic field strength) × area

  $\emptyset = B \times A$
- The unit for magnetic flux is T m^2 because B has the unit T and A has the unit m^2.
- Faraday's law of induction states that the size of the induced e.m.f. in volts is proportional to the rate at which the conductor is cutting through the flux lines.
- Lenz's law states that the direction of the induced current is such as to oppose the change that is causing it.
- You find the direction of induced currents, given the direction of motion of the conductor and the direction of a magnetic field, using the right-hand rule.
- The factors that affect the magnitude of induced e.m.f. in a conductor are the speed at which the conductor cuts the lines of flux and the strength of the magnetic field.
- Inductance is defined as the property in an electrical circuit where a change in the electric current through that circuit induces an e.m.f. that opposes the change in current. The Henry is its SI unit.
- Self-inductance is the inductance that arises within a coil. Mutual inductance occurs when a coil is in close proximity to another.
- A diagram of a transformer is as shown in Figure 4.68.

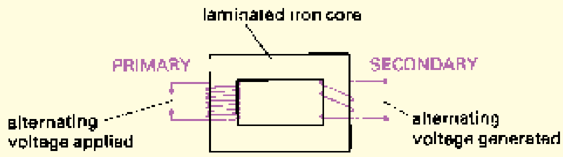


Figure 4.68

- For an ideal transformer, $P_{out} = P_{in}$

  For an ideal transformer, $\frac{V_s}{V_p} = \frac{N_s}{N_p} = \frac{N_s}{N_p}$

Review questions

1. Define magnetic flux and give its SI unit.
2. State Faraday's law of induction.
3. Describe a simple experiment that demonstrates an induced e.m.f. caused by changing magnetic flux.
4. State Lenz's law.
5. Describe how you could find the direction of induced currents, given the direction of motion of the conductor and the direction of a magnetic field.
6. State the factors that affect the magnitude of induced e.m.f. in a conductor.
7. Describe the link between electricity and magnetism.
8. What law can be used to calculate the magnitude of induced e.m.f.?
9. Define inductance and give its SI unit.
10. What are the differences between self-inductance and mutual inductance?
11. Describe and explain the action of the simple a.c. generator.
12. Draw up a table to compare the actions of d.c. and a.c. generators.
13. Draw a diagram of a transformer.
14. Give a simple explanation of the principles on which a transformer operates.
15. State the ideal transformer equations.

End of unit questions

1. Why would you use a magnetic shield?
2. Explain why some materials are better magnetic shields than others.
3. Explain the magnetisation of a nail in terms of what happens to the domains.
4. a) What is the formula for finding the magnetic field strength at a point due to a current-carrying wire?

   b) Calculate the magnetic field strength when a current of 3 A flows in a wire 2.5 m long and produces a force of 15 N.
5. What rule do you apply to find the direction of magnetic field lines around a straight current-carrying wire?
6. Calculate the force felt by a 20 cm wire carrying 1.5 A at right angles to a magnetic field of 6×10^{-5} T.
7. What information do you need to find the magnetic field strength in a solenoid?
8. Explain why the magnetic field in a solenoid is greater if the solenoid has an iron core.
9. What is the relationship between magnetic field strength, mass of particle, velocity of particle, charge on particle and radius of path?
10. In a machine called a mass spectrometer, charged particles follow a curved path. For these particles

    $Bqv = \frac{mv^2}{R}$ where R is the radius of the path.

    a) Show that this rearranges to $\frac{q}{m} = \frac{v}{BR}$.

    The particles are accelerated to speed v by an electric field and their kinetic energy $\frac{1}{2}mv^2 = qV$

    b) Show that this gives $v = \sqrt{\frac{2qV}{m}}$

    c) By substituting this into the expression for $\frac{q}{m}$, show that $\frac{q}{m} = \frac{2V}{B^2 R^2}$

    d) Find the radius of the path of $^{35}Cl^-$ ions in such a machine if $V = 3.0$ kV and $B = 3.0$ T. ($q = 1.67 \times 10^{-27}$C)
11. What is the motor effect?
12. State Faraday's law of induction.
13. What is the induced e.m.f. when the rate of change in magnetic flux is 6 T m^2 s^{-1}?
14. What is the induced e.m.f. when a current of 3 A flows for 2 s in an inductor of inductance 3×10^{-3} H?
15. What is the ideal transformer equation?
16. a) Why will a transformer work only from an alternating supply?

    b) What would be the effect of connecting it to a battery instead?

Introduction to electronics Unit 5

Contents

Electronics is the study and design of systems that use the flow of electrons through such components as semiconductors, resistors and capacitors. Many of the concepts which we met in Unit 3 (Current Electricity) are also relevant to electronic circuits, but several of the components used in electronic circuits have very specialised characteristics.

5.1 Vacuum tube devices

By the end of this section you should be able to:

- Define the term electronics.
- State the importance of electronics in your daily life.
- State what is meant by thermionic emission.
- Describe the behaviour of vacuum tubes.
- Describe the function of a cathode ray tube.
- Describe the uses of a cathode ray tube.
- Represent both d.c. and a.c. on current–time or voltage–time graphs.
- Use the current–time or voltage–time graphs to find the period and frequency of alternating currents or voltages.

Activity 5.1: Observing objects

Observe the objects around you – at home, at school, when travelling – and note down how many are dependent on electronics.

At school, display the objects you have considered on a concept map and discuss them in the class.

KEY WORDS

thermionic emission *the escape of electrons from a heated metal surface*

cathode ray oscilloscope *electronic test equipment that provides visual images of electrical signals and oscillations*

Thermionic emission

In Unit 3 we learned about electrons (conduction electrons) that are free to move around within a metal at random, even without the application of a voltage to cause a general drift in one direction. If the metal is heated up, they move faster, and some of the more energetic electrons can then escape from the surface of the metal in a way very similar to evaporation of the faster molecules from the surface of a liquid.

The effect is known as **thermionic emission**. Its rate is negligible at ordinary temperatures, and does not become significant until high temperatures are reached.

Thermionic emission provides a controllable supply of electrons in a vacuum. This is the basis of the **cathode ray oscilloscope**, a piece of equipment that has important uses in studying and displaying electrical signals or oscillations.

Vacuum tubes

The thermionic diode

The thermionic diode, a vacuum tube, is now almost obsolete, its job is now being performed by the semiconductor diode, which we shall study later in this unit. Nevertheless it is still useful to take a quick look at how it worked. The name **diode** refers to the fact that the device has two electrodes – an anode and a cathode. A metal plate (the cathode of Figure 5.1) was heated so as to emit electrons. This was done electrically by placing the cathode in front of what amounted to an electric heater, as shown. This heater was a tungsten wire like the filament of a light bulb, being run at red heat. It operated off a few volts, either a.c or d.c. – the 6 V a.c. in the drawing was typical.

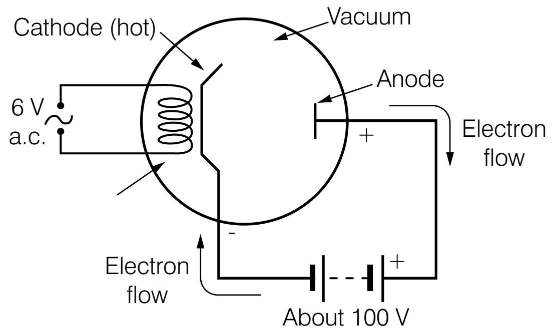


Figure 5.1 *Thermionic diode.*

The electrons 'boiled off' from the hot cathode and were attracted to the anode – a cold metal plate which was commonly at about +100 V with respect to the cathode. In this way the circuit was completed, and current flowed.

The main use of the device was for **rectification** – to obtain a d.c. current from an alternating voltage. It behaved like a valve in a water pipe and permitted one-way flow only. Current could pass across it in the direction shown in Figure 5.1 but if the polarity of the main battery was reversed, the cold anode would not emit electrons into the vacuum and so current would not flow. A diode valve was wasteful of power because of the heater.

Cathode rays

Cathode rays are a beam of electrons moving through a vacuum at a high speed. They are produced by an **electron gun**, which is a vacuum tube device. A simple example of one is shown in Figure 5.2. Practical examples may include arrangements for controlling the number of electrons in the beam and for focusing the beam to counteract the natural tendency of the particles to spread out. Nevertheless the drawing shows the principles.

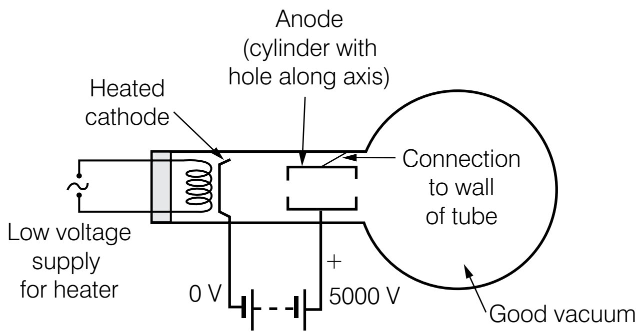


Figure 5.2 *Electron gun.*

KEY WORDS

diode *an electrical component with two electrodes, used for rectification*

rectification *converting alternating current to direct current*

electron gun *an electrical component producing a beam of electrons moving through a vacuum at high speed*

As in the thermionic diode, electrons are released from the heated cathode by thermionic emission, and are attracted towards the anode.

Unlike the diode, the voltage on the anode may be as high as 5000 V relative to the cathode. The electrons are pulled towards the centre of the anode, accelerating to a very high speed – of the order of a tenth that of light – shooting straight through the hole along the axis of the anode.

In the early days

Early workers in this area of science were puzzled by this strange radiation which could pass through wood and flesh and cause a photographic plate to blacken; they were unaware of its dangers, and did not realise that it was a previously undetected part of the family of electromagnetic waves (see Unit 6). The name they chose – X-rays – indicates clearly the extent of their knowledge at the time!

X-ray tube

The X-ray machine (Figure 5.3) is also a vacuum tube device. Electrons are released from the cathode by thermionic emission and are then accelerated through a p.d. of the order of 100 kV so as to hit the anode at an extremely high speed. When such fast-moving cathode rays are suddenly stopped, X-rays are produced.

Most of the beam's energy is released as heat rather than X-rays, so the anode gets very hot. To minimise this it is made of a large block of copper to conduct the heat away – the other end being equipped with cooling fins or having cold liquid pumped round it. Even so, the working tip of the anode is made from tungsten because of its high melting point.

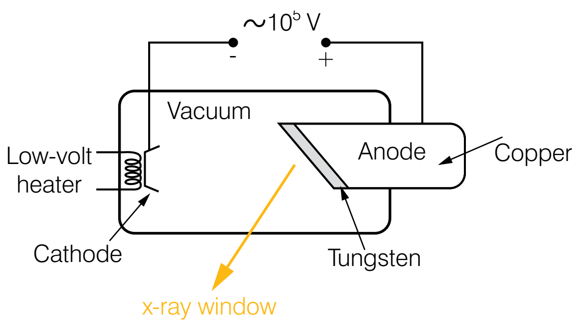


Figure 5.3 *X-ray tube.*

Cathode ray oscilloscope (CRO)

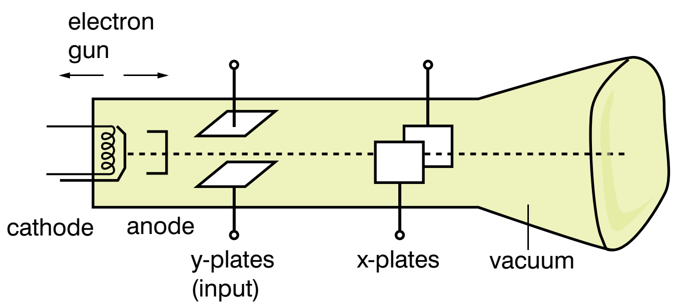


Figure 5.4 *The cathode ray oscilloscope.*

Another very useful vacuum tube device is the cathode ray oscilloscope (Figure 5.4).

In a cathode ray oscilloscope the beam of electrons produces a spot on a fluorescent screen at the end of the tube. Before reaching the screen it passes between two sets of deflecting plates – one pair to deflect it in the y-direction and the other in the x-direction.

Any rapidly varying voltage you wish to examine is connected across the y-plates. This may be the output from a microphone, for instance, or from an a.c. power supply. This voltage will cause the spot to move quickly up and down the screen (the deflection being proportional to the voltage), so a vertical line is visible. Because of the very low mass of the electrons which make up the beam, they can respond to very rapidly changing voltages.

KEY WORDS

direct current (d.c.) *an electric current that flows in a constant direction*

gain control *a device adjusting the amount of beam deflection in a cathode ray oscilloscope*

Some uses of the CRO

Direct current

The CRO can be used as a voltmeter and will represent the voltage of a source of **direct current** as a stationary spot of light on the screen. It can be calibrated by using a known voltage across the y plates so that the value of an unknown voltage can be measured.

Sensitivity

The sensitivity – the size of the deflection caused by the voltage applied across the y plates of a CRO – can be adjusted using the **gain control**. For example, if the sensitivity is set to 3 V per cm and the spot is deflected by 2 cm by an unknown p.d., the value of the p.d. must be 2 × 3 = 6 volts.

Worked example 5.1

A 1.5 V cell is connected to the y-plates of a CRO and the gain control adjusted so that the trace is 1 cm above the zero line (Figure 5.5a). The cell is removed, an unknown p.d. applied to the y-plates and a new trace is seen on the screen of the CRO (Figure 5.5b). What is the size of the unknown p.d.?

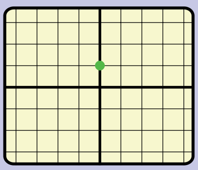


Figure 5.5a

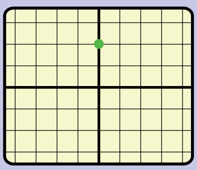


Figure 5.5b

A 1.5 V cell is represented by a deflection of 1 cm. The unknown p.d. is represented in Figure 5.4b by a 2 cm deflection. The unknown p.d. is therefore 3 V.

Time base

To discover more about a signal being studied a voltage called a time base is applied to the x-plates. This pulls the spot across the screen from left to right at a steady rate; it then flies back to the left-hand side and repeats its movement. If the time base is applied to a signal made by a direct current, the trace changes from a dot (Figure 5.6a) to a line (Figure 5.6b).

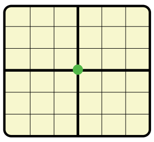


Figure 5.6a *Time base off.*

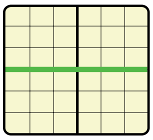


Figure 5.6b *Time base on.*

The time base control can be calibrated (in milliseconds per cm) to show how long the spot takes to cross each centimetre of the screen. For example, if it is set to 1 ms/cm the spot takes 1 ms (1 millisecond) to move 1 cm to the right. This can be very useful in measuring distances if, for example, traces of a sonar signal and its echo are studied.

Provided the gain control and the time base are synchronised (and the instrument does this for you), what appears on the screen is an apparently static graph of the input voltage plotted against time.

Alternating current

The voltage of an alternating current varies between a (positive) maximum to a (negative) minimum. If the time base is switched off, this is represented as a vertical line on the screen of a CRO (Figure 5.7a). If the time base is switched on, this becomes a curve (Figure 5.7b) whose shape is known as a **sine wave**.

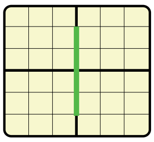


Figure 5.7a *Time base off.*

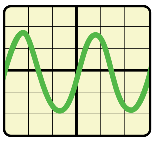


Figure 5.7b *Time base on.*

KEY WORDS

sine wave *a mathematical function that describes a smooth repetitive oscillation*

Finding the period and frequency of alternating currents or voltages using the CRO

In the same way that an unknown d.c. voltage can be measured using a CRO by comparing its trace with that made by a known p.d., the CRO can measure frequency by comparing a wave of unknown frequency with one of known frequency.

The known signal is applied to the CRO and the time base adjusted so that one complete wave appears on the screen (Figure 5.8a). The unknown signal is then applied in place of the known signal, without altering any of the CRO controls, and the trace on the screen is studied (Figure 5.8b).

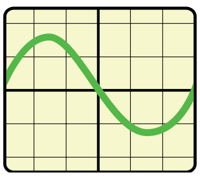


Figure 5.8a *(50 Hz).*

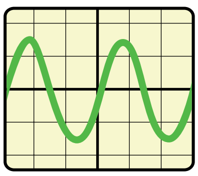


Figure 5.8b *(unknown).*

As Figure 5.8b shows two complete waves it follows that the unknown frequency is twice that of the known signal, i.e. $2 \times 50 = 100$ Hz.

Worked example 5.2

Figure 5.9a shows the trace on a CRO made by a 100 Hz signal. What is the frequency of the signal in Figure 5.9b (which was made without altering any of the CRO controls after measuring the 100 Hz signal)?

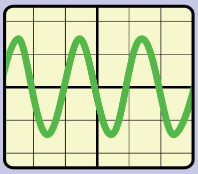


Figure 5.9a

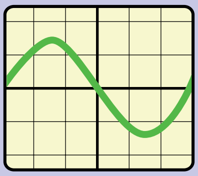


Figure 5.9b

The 100 Hz signal in Figure 5.8a shows three complete waves while the signal of unknown frequency in Figure 5.8b shows one wave. The unknown frequency in this example is therefore one-third that of the known signal

The unknown frequency = $100 \times \frac{1}{3} = 33.3$ Hz

Worked example 5.3

The sensitivity of a CRO is set as 5 V/cm, and the time base setting is 1 ms/cm. Find the peak voltage and frequency of the a.c. signal shown in Figure 5.10.

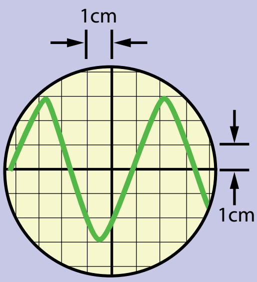


Figure 5.10

The maximum vertical displacement is 3 cm. Each centimetre of vertical displacement corresponds to 5 V.

The peak voltage of the signal shown in Figure 5.10 is therefore $5 \times 3 = 15$ V.

The time base setting is 1 ms /cm. The spot therefore takes 1 ms to move horizontally by one centimetre. One complete cycle of the signal shown in Figure 5.10 takes 5 cm. The time for one cycle is therefore 5 ms = 5×10^{-3} and there must therefore be 200 cycles each second.

The frequency of the a.c. signal is therefore 200 Hz.

TV picture tube

The older sort of television, with the big heavy tube, is similar to the CRO. The receiver sends currents through coils that are mounted just outside the tube; their magnetic fields deflect the spot of light on the screen so it moves rapidly to trace an ordered path all over the screen.

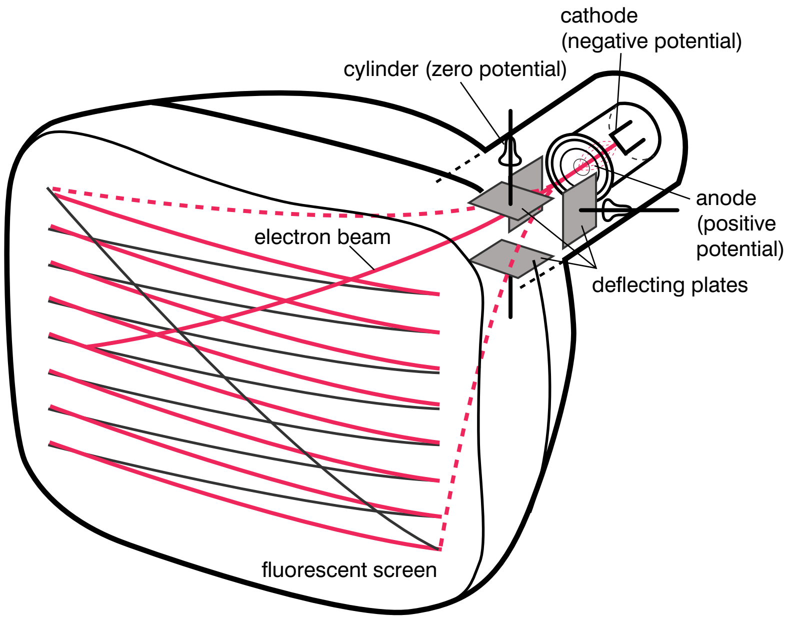


Figure 5.11a *Television tube.*

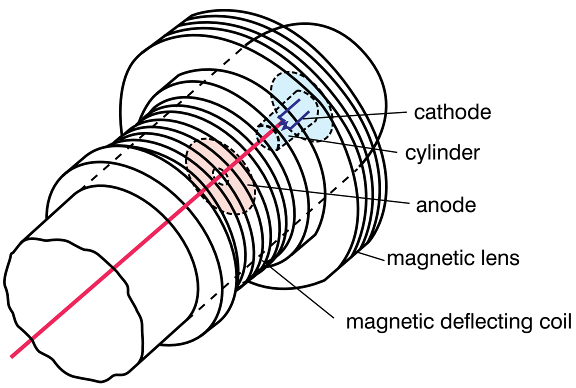


Figure 5.11b *Coils forming a magnetic lens.*

Summary

- Thermionic emission is the escape of conduction electrons from a hot metal surface.
- Thermionic electrons may be accelerated through a high voltage to produce a beam of cathode rays.
- These cathode rays convey negative charge, and may be deflected accordingly by magnetic and electric fields.

Review questions

1. In an oscilloscope tube what is the purpose of:
   a) the heater
   b) the cathode
   c) the anode
   d) the x- and y-plates?
2. Explain what it means if the time base of an oscilloscope is set at 2 ms/cm (1 ms = 0.001 s.)
3. Figure 5.12**a** shows a reading of 1.5 V. Give the values displayed in **b**, **c** and **d**.

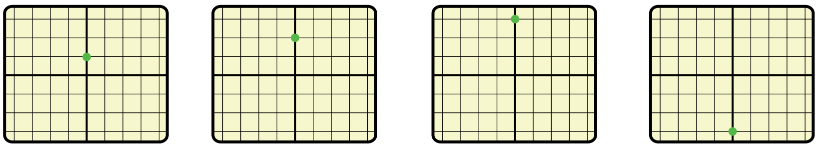


Figure 5.12a *Figure 5.12b* *Figure 5.12c* *Figure 5.12d*

4. Explain carefully why a cathode ray oscilloscope:
   a) acts as a voltmeter
   b) responds to changes in that voltage within less than a millionth of a second
   c) draws virtually no current from the voltage source.
5. Why is the anode of the X-ray tube drawn in Figure 5.3 made of copper with a tungsten tip?
6. Describe three different uses for X-rays. Say what precautions should be taken while using them, and explain why.

5.2 Conductors, semiconductors and insulators

By the end of this section you should be able to:

- Distinguish between conductors, semiconductors and insulators.
- Give examples of semiconductor elements.
- Distinguish between instrinsic and extrinsic semiconductors.
- Describe a semiconductor in terms of charge carriers and resistance.

Conductors, semiconductors and insulators

We saw in Unit 3 that materials can be divided into three classes:

- Insulators, such as glass and plastic, which do not conduct electricity because every electron in them is tightly bound to its parent atom.

KEY WORDS

intrinsic semiconductor *a pure semiconductor not containing any dopant*

hole *the lack of an electron at a position where one could exist in an atomic lattice*

intrinsic conduction *electrons and holes in a semiconductor moving in opposite directions when an e.m.f. is applied*

- Conductors such as metals. All the electrons in the inner shells are still tightly bound to their atoms. However, those electrons in the outermost shell of every atom are free to move within the metal. We describe them as being in the conduction band.
- Semiconductors. Just a few of the outermost electrons have enough energy to be in the conduction band (that is, to break free from their parent atom), but this number rises as the material becomes hotter. Silicon, germanium, lead sulphide, selenium and gallium arsenide are all semiconductors.

Intrinsic semiconductors

Conduction in a pure (intrinsic) semiconductor

Substances such as silicon and germanium have resistivities between those of insulators and those of conductors. These substances are known as semiconductors, and they form the basis of many of the devices that we take for granted in a technological society. Pure semiconductors are usually referred to as **intrinsic semiconductors**, since their conductivity is not affected by any external factors. As we shall see in Section 5.3, trace impurities can greatly alter the conductivity of semiconductors.

In an intrinsic semiconductor, electric current is carried by moving electrons, as it is in metals, although the number of charge carriers in silicon is perhaps a billion times fewer than in copper. However, in addition to electrons, intrinsic semiconductors can also be considered to contain moving positive charges that carry current. This can be explained as follows.

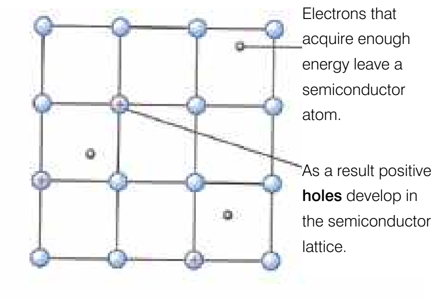


Figure 5.13a

Lattice structure of the atoms in an intrinsic semiconductor

Some of the electrons in an atom of an intrinsic semiconductor are held less tightly than others. This means that in a piece of intrinsic semiconductor material at room temperature there will always be a few free electrons that have been ‘shaken free’ of their atoms by thermal excitation (when the material has absorbed energy from the surroundings). When an electron leaves an atom in this way, the atom becomes positively charged (Figure 5.13a). The effect of an electron leaving an atom is therefore to create a positive charge in the semiconductor lattice. This positive charge is called a **hole**.

When an electric field is applied to the semiconductor (that is, when it is connected to a source of e.m.f.) the electrons and holes move in opposite directions, and the semiconductor exhibits **intrinsic conduction**.

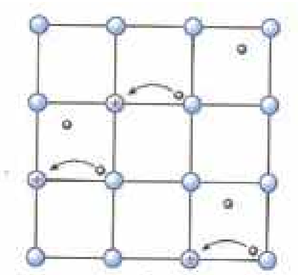


Figure 5.13b

This happens because, under the influence of this electric field, electrons still bound to atoms in the lattice are able to move through the lattice from an atom to a nearby hole (Figure 5.13b), thus causing the hole to appear to move through the lattice (Figure 5.13c). This motion happens in the opposite direction to the motion of the electrons.

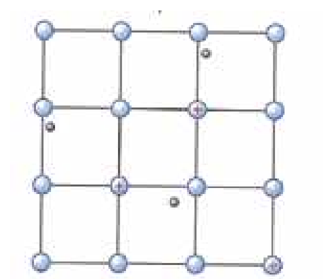


Figure 5.13c

The current in a pure semiconductor consists of free electrons moving through the semiconductor lattice in one direction, with an equal number of positively charged holes moving in the other direction (see Figure 5.14).

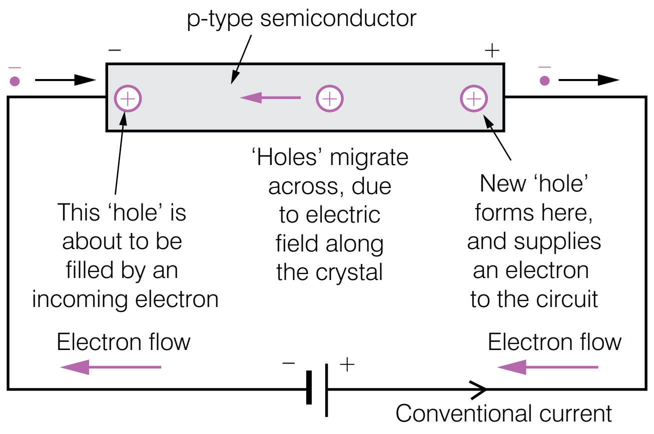


Figure 5.14

Charge carriers and resistance in a semiconductor

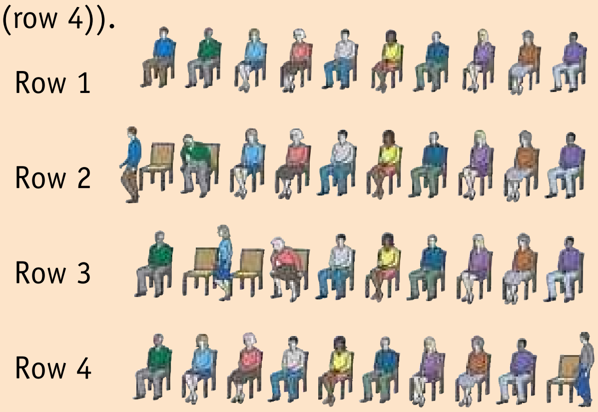
While the resistance of metallic conductors rises as they warm up, with semiconductors the resistance falls greatly as their temperature goes up.

This is because when the temperature of a semiconductor is raised, more electrons (charge carriers) have enough energy to break free. As the number of charge carriers increases, the resistance of the semiconductor material decreases and the material conducts better.

Activity 5.2: Mysteriously moving chair

Place 10 chairs in a row and ask 10 people to sit on the chairs (Figure 5.15 row 1). Ask the person at one end of the row to stand up and move away (Figure 5.15 row 2). Ask the person sitting next to the empty chair to move into that chair and then ask the remaining people to move to the empty chair as their neighbour moves up (row 3). Eventually the empty chair will be at the end of the row (a new person could take the empty chair! (row 4)).

Figure 5.15



The motion of holes through the lattice is like the motion of the empty chair in this line. As the people move from right to left, the empty chair moves from left to right.

Summary

- In a semiconductor just a few electrons have enough energy to break free from their atoms. These enable it to carry a current.
- As the electrons move in one direction, positive 'holes' appear to move in the other direction.

Review questions

1. Explain why a semiconductor can conduct at all.
2. Explain why a semiconductor conducts better when it is hot.
3. Explain what we mean by a positive 'hole' in a silicon crystal.

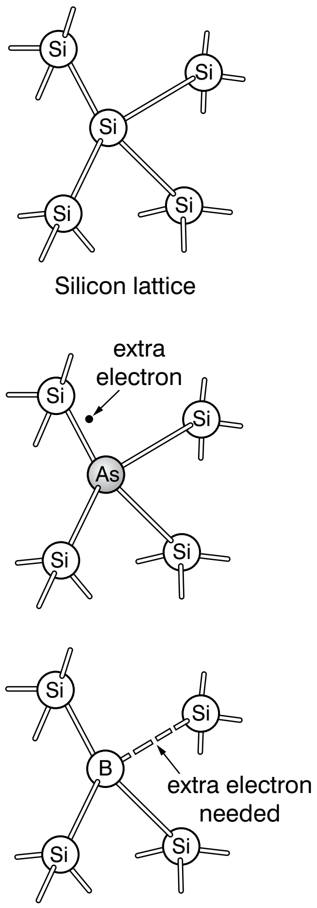


Figure 5.16 *Dopants in a silicon lattice.*

5.3 Semiconductors (impurities, doping)

By the end of this section you should be able to:

- Explain doping to produce the two types of semiconductors.
- Indentify semiconductors as p-type and n-type.
- Describe the mode of conduction by the majority and minority carriers.
- Define the term diode and show its circuit symbol.
- Draw a current versus voltage characteristics (graph) to show the behaviour of p-n junction.
- Describe how a semiconductor diode can be used in a half-wave rectification.
- Sketch voltage–time graphs to compute the variation of voltage with time before and after rectification.
- Distinguish between direct current from batteries and rectified alternating current by consideration of their voltage–time graphs.
- Show the circuit symbols of semiconductor devices such as thermistor, LED, LDR and transistors.

KEY WORDS

doping *deliberately introducing impurities into a semiconductor to change its electrical properties*

extrinsic semiconductor *a semiconductor that has been doped*

Doping

The variety of uses to which semiconductor materials such as silicon can be put in devices such as transistors and diodes depends on introducing minute quantities of impurities into their atomic lattice structure in a process called **doping**. This process introduces extra charge carriers by replacing atoms in the semiconductor lattice with atoms of an impurity of similar size (this is important so as to minimise the distortion of the semiconductor lattice). This alters the semiconductor's conducting properties by introducing extra large carriers to the semiconductor lattice, forming what is called an **extrinsic semiconductor**.

Majority and minority carriers

Silicon is in group 4 of the Periodic Table, which means the outer shell of the neutral silicon atom contains four electrons.

Suppose that atoms of an impurity from group 5 of the Periodic Table (neutral atoms in group 5 have five electrons in their outer shell) such as arsenic are added to the silicon lattice. The crystal as a whole remains uncharged, but these impurities help to provide 'spare' free electrons to the crystal, causing much improved conduction. Since the **majority carriers** in this type of semiconductor material are negative electrons, we describe this as an **n-type semiconductor**. (The **minority carriers** in n-type semiconductor material are holes.) Arsenic is described as a **donor** impurity, because it releases free electrons into the lattice.

Alternatively, if silicon is doped with impurity atoms from group 3 of the Periodic Table (neutral atoms in group 3 have three electrons in their outer shell) such as boron, there is again no overall charge. However, the apparent electron deficiency can shift from one atom to another and so behave like a kind of positive 'hole' which can move through the crystal and thereby carry a current through it. These 'holes' are the majority carriers in this type of semiconductor – called **p-type semiconductor** – and the minority carriers are electrons. Boron is an example of an **acceptor** impurity which traps electrons when introduced into the lattice, resulting in an increase in the number of positive holes.

Conduction in a doped semiconductor

The number of free electrons and holes can be altered dramatically by doping. For example, the addition of only one arsenic atom per million silicon atoms increases the conductivity 100 000 times.

Figure 5.17 shows how these processes occur .

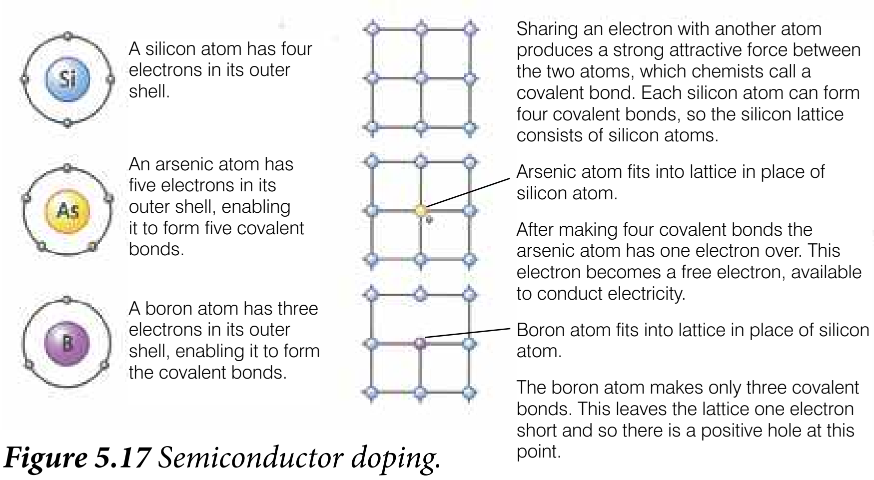


***Figure 5.17** Semiconductor doping.*

KEY WORDS

majority carriers *the type of carrier, electron or hole, that constitutes more than half the carriers in a semiconductor*

n-type semiconductor *a semiconductor in which the majority carriers are electrons, due to doping*

minority carriers *the type of carrier, electron or hole, that constitutes less than half the carriers in a semiconductor*

donor *impurity atoms added to a semiconductor which release free electrons*

p-type semiconductor *a semiconductor in which the majority carriers are holes, due to doping*

acceptor *impurity atoms added to a semiconductor which trap electrons*

junction *the region where two types of semiconducting materials touch*

Activity 5.3: Human wire

In this activity a group of students will model a doped p-type semiconductor lattice – a silicon lattice with boron impurities.

In a group of nine, each student will a model silicon atom – each person will have four rocks to represent four electrons.

One student needs to represent a boron atom – s/he will have three rocks to represent electrons and one basket to represent a hole.

Students should move to show how the electron travels in one direction and the hole in the other.

Activity 5.4: Modelling the lattice of a semiconductor

Figure 5.18 shows the three-dimensional lattice structure of a semiconductor such as silicon. Each silicon atom has four electrons available for making bonds with other atoms

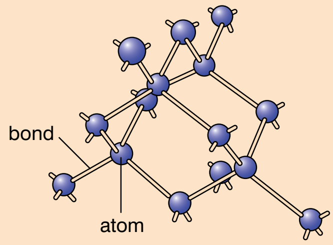


Figure 5.18 *Lattice structure.*

Using gummy sweets to represent silicon atoms and toothpicks to represent the bonds made by a silicon atom, build a lattice like that shown in Figure 5.18, where each silicon atom is joined to four others.

Change your lattice to show how doping affects the structure.

To represent an n-type semiconductor, stick an extra toothpick into some of the gummy sweets. A sweet with five toothpicks represents a donor impurity (such as arsenic) because the fifth electron is free to move into the lattice.

To represent a p-type semiconductor, remove one of the gummy sweets in the lattice and replace it with a marshmallow. The marshmallow represents an acceptor impurity (such as boron) – think of the squashy marshmallow as representing a 'hole' into which electrons would be attracted.

The p–n junction diode

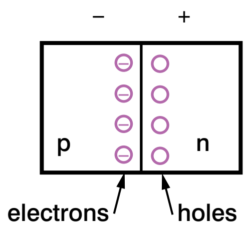


Figure 5.19 *p–n junction.*

Suppose a p-type semiconductor (with 'spare' positive holes) is in contact with an n-type one (with 'spare' mobile electrons). It is important to realise that before they were brought into contact both pieces of semiconductor were electrically neutral overall. At the junction where they meet, some of the n-type's electrons move, or 'fall' into the p-type's holes (Figure 5.19). This movement is known as idffusion current.

This diffusion current causes the p-type to become slightly negative while the n-type is left equally positive, leaving a 'depletion zone' for a small distance each side of the boundary (of the order of 1 μm) – a shortage of 'holes' on one side and of free electrons the other (Figure 5.20).

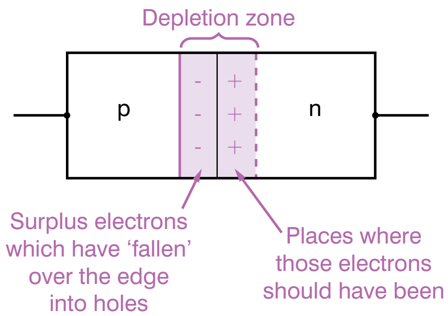


Figure 5.20 *A p–n junction.*

This adverse voltage (of about 0.6 V for silicon) between the p-type and n-type semiconductors will prevent any more electrons from crossing the boundary, so the diode will not conduct in that direction. In the depletion zone there are no more 'holes' in the p-type and no free electrons in the n-type, so it forms a non-conducting strip which blocks all current.

Forward and reverse bias

Imagine a cell connected as shown in Figure 5.21. Remember that it tries to move electrons from its negative terminal round the circuit to its positive terminal, which will only serve to make the situation at the boundary worse. After this momentary transfer, the voltage at the boundary (which is opposing that of the cell) becomes as large as that of the cell. No current will then flow. We say the junction is **reverse biased**.

If instead you apply a voltage across the diode to make the p-type positive, and make it more than 0.6 V so as to overcome the depletion-zone voltage across the junction, then the junction is said to be **forward biased** (Figure 5.22). In this direction the cell is helping to keep the transfer going. Excess electrons are being removed from the p-type, and are being fed back into the *n*-type to fill the extra 'holes'. That could go on forever – in this direction the junction will conduct.

Electrons will exit from the left-hand lead, continually leaving fresh positive 'holes' behind them. The voltage across the crystal will cause these holes to move to the right. At the junction, each hole meets a free electron and so ceases to exist. All the time, replacement electrons are flowing round the circuit and into the n-type via the right-hand lead and so the whole process is maintained.

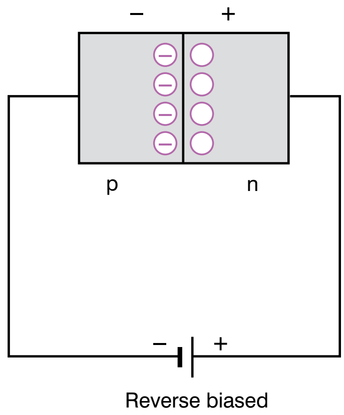


Figure 5.21 *Reverse biased, no current.*

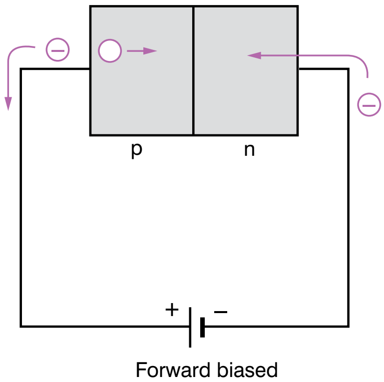


Figure 5.22 *Forward biased, current flows.*

Current–voltage characteristics of the semiconductor diode

The p–n junction as described acts as a diode: in one direction it will conduct, in the other direction it will not. The behaviour of such a diode may be illustrated by a current–voltage graph like the one shown in Figure 5.23.

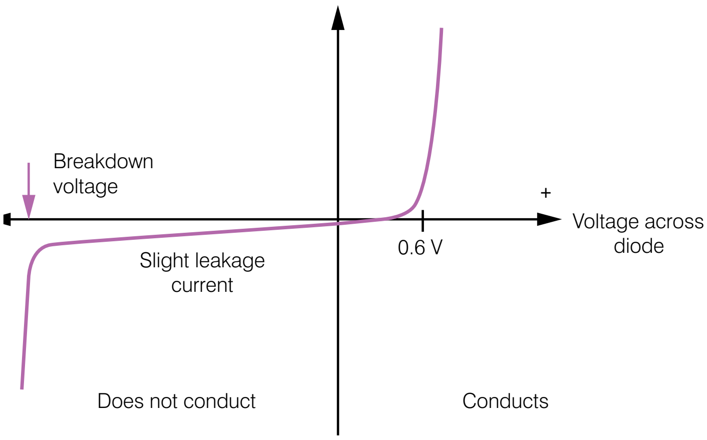


Figure 5.23

In the forward direction, silicon requires about 0.6 V before conduction will start, but after that the current is usually limited by the resistance of the rest of the circuit. In the other direction, notice there will be a tiny leakage current. If the reverse p.d. becomes too great for the device, the barrier at the junction breaks and permits a large current to flow.

KEY WORDS

reverse biased *connecting the positive terminal of a cell to the n-type region of a diode, and the negative terminal to the p-type region, preventing conduction*

forward biased *connecting the positive terminal of a cell to the p-type region of a diode, and the negative terminal to the n-type region, allowing conduction*

Demonstration: Testing the conduction of a diode

If you have the equipment, you could test the conduction of a diode by connecting a battery, a variable resistor, an ammeter and a diode in a circuit. If possible, use an LED as your diode as this should give you useful visual clues about how the experiment is progressing. You might also need a small resistor in the circuit to protect the diode. Make sure that you know in which direction the current is flowing in the circuit so that you can connect the anode and cathode of the diode correctly in the circuit. Connect a voltmeter across the diode. Use the variable resistor to change the current in the circuit and observe the voltage across the diode and the current in the circuit. You should see a series of values like those displayed in Figure 5.23.

Some semiconductor devices

Diode

A diode is an electronic component with two electrodes – an anode and a cathode – which will only allow electric current to pass through it in one direction (Figure 5.24).

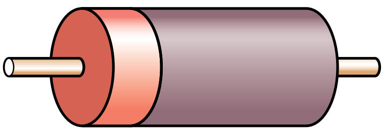


Figure 5.24a *An example of a diode (ringed end shows cathode).*

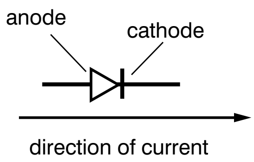


Figure 5.24b *Symbol for a diode.*

The semiconductor diode – formed from a layer of *p*-type semiconductor joined to a layer of *n*-type semiconductor material as seen above – is a very important electronic component.

LDR

A light-dependent resistor (LDR for short) conducts electricity, but in the dark it has a very high resistance. Shining light on it appears to 'unjam' it, because its resistance falls. The brighter the light, the better it conducts. The symbol for an LDR is shown in Figure 5.25b.

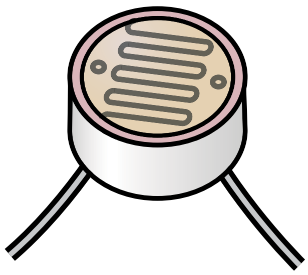


Figure 5.25a *LDR.*

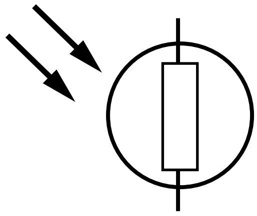


Figure 5.25b *Symbol for LDR.*

Thermistor

The thermistor shown in Figure 5.26a is a piece of semiconductor material that has a high resistance in the cold. Its resistance drops as it becomes warmer.

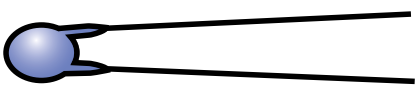


Figure 5.26a *Thermistor.*

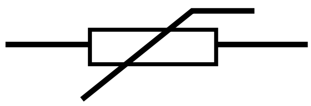


Figure 5.26b *Symbol for thermistor.*

Variable resistor

The variable resistor (Figure 5.27) is a very useful component in electronic circuits, particularly in circuits containing **transistors**.

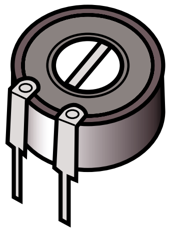


Figure 5.27a *Variable resistor.*

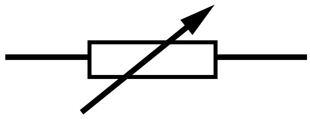


Figure 5.27b *Symbol for variable resistor.*

KEY WORDS

transistor *a semiconductor device used to amplify or switch electronic signals*

bipolar junction transistor *a device in which the current flow between two terminals, the collector and the emitter is controlled by the amount of current that flows through a third terminal, the base*

LED

The light emitting diode, whose symbol is shown in Figure 5.28b, can be seen in a multitude of devices. When a current is passed in the forward direction, an LED emits light. The LED is a very useful component – if there is one in a circuit, it is possible to see immediately if current is flowing. LEDs are now available in a range of colours – red, green, blue, white. White LEDs are increasingly being used in lighting; they produce light very efficiently (using relatively little energy).

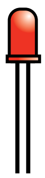


Figure 5.28a *LED.*

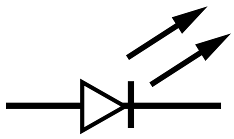


Figure 5.28b *Symbol for LED.*

Transistor

The transistor (Figure 5.29) is a very significant semiconductor component which we shall learn more of in Section 5.4.

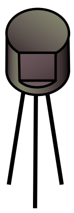


Figure 5.29a *Transistor.*

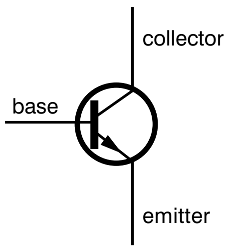


Figure 5.29b *Symbol for transistor.*

A **bipolar junction transistor** is made of three layers of doped semiconductor and it has three terminals – the base is connected to the central layer, the other two (the collector and the emitter) are each connected to one of the outer layers. Figure 5.29 shows

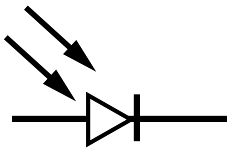


Figure 5.30 Symbol for photodiode.

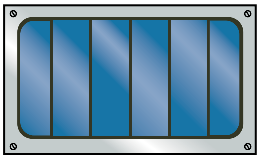


Figure 5.31a *Photovoltaic cell .*

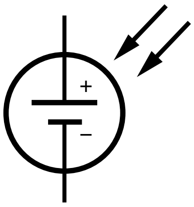


Figure 5.31b *Symbol for photovoltaic cell .*

KEY WORDS

photovoltaic cell *a cell that converts solar energy into electrical energy*

an n–p–n transistor, which has a layer of p-type semiconductor sandwiched between two layers of n-type.

Photodiode

The photodiode (Figure 5.30) is a light-sensitive diode used to detect light or to measure its intensity. Photodiodes are reverse-biased so they do not conduct. Light incident on the photodiode frees a few more electrons and the device starts to conduct.

Photovoltaic cell

The photovoltaic cell (Figure 5.31) is a form of photodiode.

The base layer of a photovoltaic solar cell is made of p-type semiconductor material. This is covered with a layer of n-type semiconductor material. When light strikes the junction between n- and p-types of semiconductor, electrons flow through the structure of the cell.

Activity 5.5: Light into power

In this experiment we study the way in which photovoltaic solar cells convert light energy to electrical energy.

When light strikes the junction between n- and p-types (see Figure 5.32), electrons flow through the crystal structure and round the circuit connected between them. This electric current can be used directly or stored in a rechargeable cell.

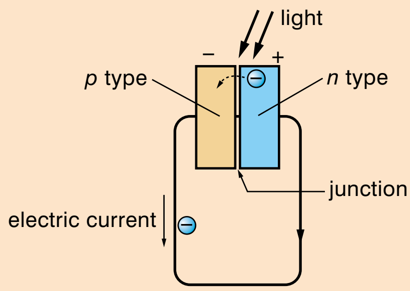


Figure 5.32

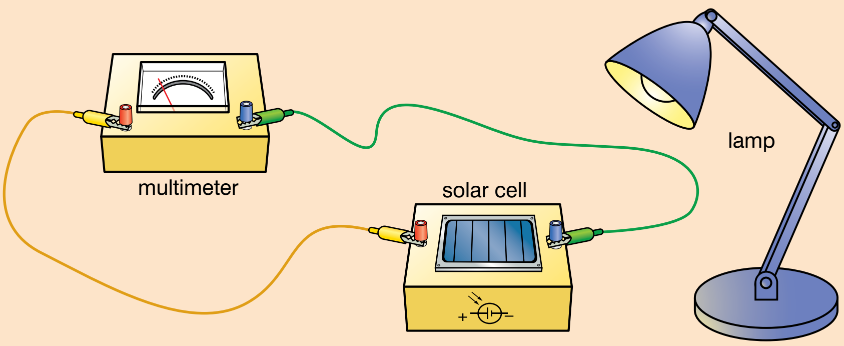


Figure 5.33 *Testing a photovoltaic cell in different light conditions.*

Connect one end of a crocodile clip lead to the positive terminal on the back of a small photovoltaic cell (0.4 V 100 mA) and the other end of this lead to the positive probe of a multimeter (see Figure 5.33). Connect the negative terminal on the back of the cell to the negative probe of the multimeter in the same way.

Switch the multimeter to read voltage – in hundredths of volts.

Cover the photovoltaic cell and record the multimeter reading in the table below (in the row 'no light, covered').

Uncover the photovoltaic cell and record the multimeter reading (observe the speed with which the multimeter reading changes when you uncover and cover the solar cell).

Move the photovoltaic cell into the sunlight and repeat the reading for sunlight (note on the table if the sun is falling directly on the solar cell).

Light source	Voltage reading
No light (covered)	
Room illumination	
Sunlight	

You will notice that the voltage is higher when more light falls on the photovoltaic cell.

You will probably see that the photovoltaic cell responds rapidly to the presence and absence of light (the voltage reading on the multimeter rises and falls abruptly).

Photovoltaic cells tend to be black because black objects are more efficient at absorbing radiation.

In order to produce a power supply for a particular appliance, a number of cells can be connected to produce the required voltage and current.

Rectification using the p–n junction diode

Using one diode

Direct current can be obtained from an alternating current generator by putting a diode in the circuit (Figure 5.34).

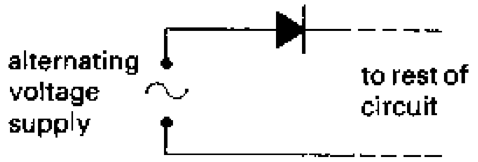


Figure 5.34 *A diode in a circuit with a supply of alternating current.*

KEY WORDS

capacitor *a device for storing electric charge, consisting of two conductors separated by a dielectric*

The diode allows the current to flow one way, but on the other half of the cycle the current cannot flow back again through the diode. The resulting current is shown in graph form in Figure 5.35. We call this half-wave rectification.

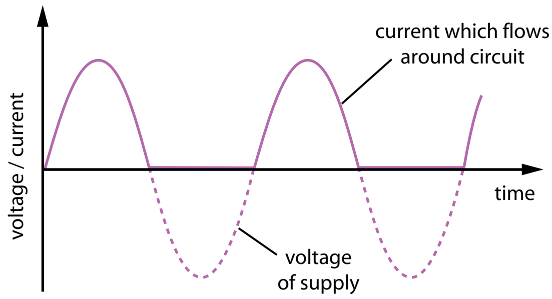


Figure 5.35 *Graph showing current flowing round circuit.*

It is direct current, because the charges always flow in the same direction and so will make progress round the circuit. However, it is not a steady smooth flow of direct current such as that obtained from a battery. Instead the charges move forward in a series of 'spurts'.

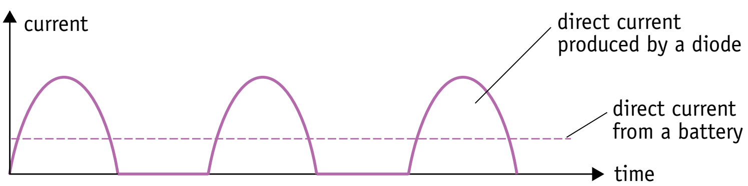


Figure 5.36 *Irregular direct current produced by diode compared with current from battery.*

Using a diode and a capacitor

It is possible to use a **capacitor** to help to smooth the fluctuations in this current. As you saw in Unit 2, a capacitor stores charge, and can release it later. The capacitor is connected across the terminals inside the casing of the power supply (Figure 5.37).

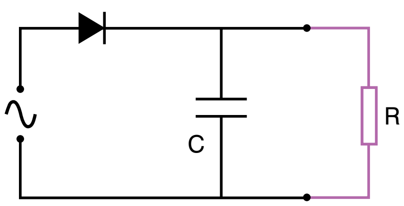


Figure 5.37

The capacitor is filled up to the peak voltage of the supply, then as the power supply's voltage drops to zero we can think of the capacitor as feeding the outside circuit. It has to keep doing this until the supply voltage next peaks and the capacitor is once more filled up (Figure 5.38).

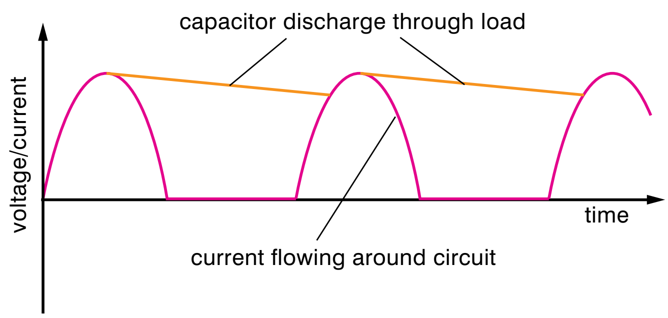


Figure 5.38

If the supply is a 50 Hz one, this will happen every $\frac{1}{50}$ s (0.02 s). The capacitor will be effective only if the time constant CR is large compared to 0.02 s, which will mean that it has emptied very little over that length of time. We often use an electrolytic capacitor as the smoothing capacitor, sometimes 1000 μF or even larger. An electrolytic capacitor is rather leaky, but since it gets refilled 50 times a second that hardly matters.

Full wave rectification

An arrangement made from four diodes is known as a **bridge rectifier**. To understand how it works needs some thought from you, so you are invited to work through it yourself. In Figure 5.39, the dotted line is the outside casing of a d.c. power pack. Inside are an a.c. power supply marked AB (obtained from the mains via a transformer) and four diodes labelled C, D, E and F. G and H are the terminals that connect it to an outside circuit, represented by resistor R.

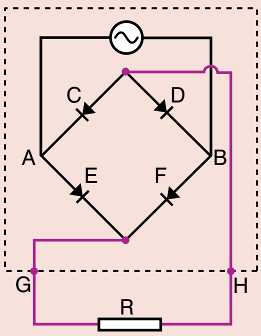


Figure 5.39

Remember that a current can flow only if a continuous circuit is provided from one terminal of power supply AB right round to the other. Check that at a moment when A is positive and B negative, conventional current has such a route through AEGRHDB. Now find the complete circuit half a cycle later, when B is the positive terminal and the conventional current starts from there. What do you notice about the direction in which the current passes through R on each half of the alternating cycle? This is known as full-wave rectification, and has the form shown in Figure 5.40.

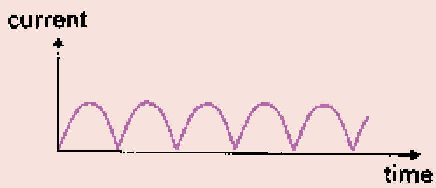


Figure 5.40

Summary

- An extrinsic conductor has its crystal doped with just a few other atoms, which can help it to conduct still better.
- A p–n junction acts as a diode, and will conduct in one direction only.
- A semiconductor diode can produce half-wave rectification from an alternating supply.

KEY WORDS

bridge rectifier *an arrangement of four diodes which produce full-wave rectification of an alternating current*

Review questions

1. An arsenic atom is about the same size as a silicon atom. Silicon has four electrons in its outer shell, arsenic has five.
   a) Explain how these two facts enable us to dope a silicon crystal with small numbers of arsenic atoms, and this makes it conduct much better.
   b) Why do we call it an n-type semiconductor?

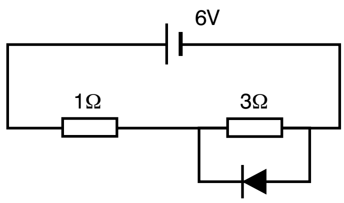


Figure 5.41

2. a) What size current will be drawn from the battery in Figure 5.41? Give your reasoning.

   b) If the battery is turned round, what size current will it now supply? Explain.

   (Assume the diode is an ideal one, so it requires zero voltage in the forward direction and possesses infinite resistance in the reverse direction.)

3. Explain in your own words why a diode valve acts as a rectifier.

4. A p-type material is brought into contact with an n-type material. Explain in your own words:

   a) Why both materials are initially uncharged.

   b) Why some free electrons from the n-type 'fall' over the boundary.

   c) Why a voltage appears across the junction, the p-type being negative.

   d) What we mean by a depletion zone, and why this will not conduct.

5. For your own notes make a table with four columns. They should be headed diode, light emitting diode (LED), light-dependent resistor (LDR) and thermistor. Under each heading give its circuit symbol and then give a brief description of how the component behaves when it is part of an electrical circuit.

5.4 Transistors (p–n–p, n–p–n)

By the end of this section you should be able to:

- Distinguish between p–n–p and n–p–n transistors.
- Identify the base, emitter and collector of a transistor.
- Use the following terms correctly: forward biased and reverse biased.
- Describe the behaviour of semiconductor devices such as thermistor, LED, LDR, photodiode and transistors.
- Use the circuit symbols for the gates.
- Draw the truth tables for the different logic gates and for a combination of logic gates.
- Explain the action of logic gates: NOT, OR, AND, NOR, NAND.

The bipolar junction transistor

Transistors use the input of relatively small signals to control circuits carrying large currents. This makes them very important as switches and amplifiers.

As we saw in Section 5.3, the bipolar junction transistor is a three-layer semiconductor device. The outer layers can be n-type materials with a p-type layer in the centre (in which case it is described as an n–p–n transistor), or it may be a p–n–p transistor with the layers the other way round (Figure 5.42).

A transistor of this type has three terminals – one to each semiconductor layer. The connection to the central layer is known as the **base**. The outer two are called the **collector** and the **emitter**.

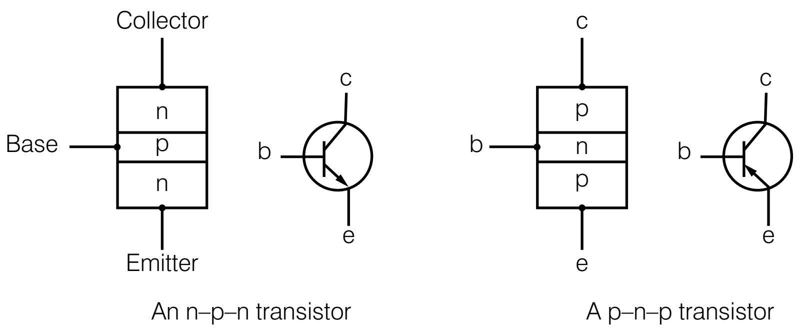


Figure 5.42 *Transistor layouts.*

The arrow on the circuit symbol indicates the way conventional current should flow through it. The term 'bipolar' means that both negative electrons and positive 'holes' play a part in conduction through the transistor, but it should be remembered that in the connecting leads the current will actually be a flow of electrons in the opposite direction to the arrow.

Such an arrangement is sometimes described as two diodes back to back. We have already seen that the diode (a single layer of n-type semiconductor joined to a layer of p-type semiconductor) only allows current to flow in one direction, when the p-type layer is connected to a positive voltage. It might therefore seem that two diodes back to back would never allow current flow, but the important thing about the transistor is that, as we shall see below, by applying a small voltage to the base connection, the transistor can be 'unblocked' and current will flow through the it. This is the feature which makes the transistor so valuable as an electrical component.

Two essential features of the transistor are:

- The base layer has to be extremely thin.
- The collector must be arranged so as to be in physical contact with and surround as much of the base as possible.

KEY WORDS

base *one of three regions forming a bipolar junction transistor. This layer separates the emitter and collector layers. If a voltage greater than around 0.6 V is applied to the base terminal, current will flow through the transistor from base to emitter*

collector *one of the three regions forming a bipolar junction transistor. When used in a circuit, a positive voltage is applied to the collector terminal.*

emitter *one of the three regions forming a bipolar junction transistor. Current will only flow from the emitter terminal if a voltage greater than around 0.6 V is applied to the base terminal.*

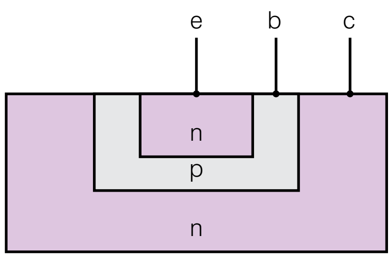


Figure 5.43 *n–p–n transistor – e (emitter), b (base) c (collector).*

Transistor biasing

'Biasing' refers to placing voltages across the terminals of a device. A diode is said to be 'forward biased' if a voltage is applied to it which enables it to conduct (from Figure 5.44 that will be when the positive side of the battery is connected to the p-type material). If the voltage is applied the other way round, the diode is said to be 'reverse biased' and it will not conduct.

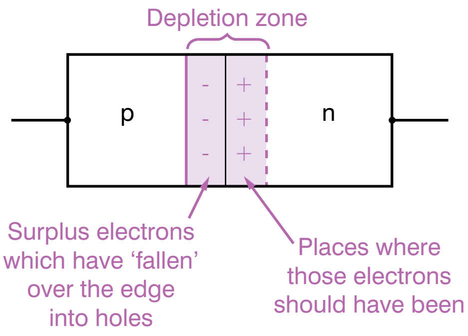


Figure 5.44 *A p–n junction.*

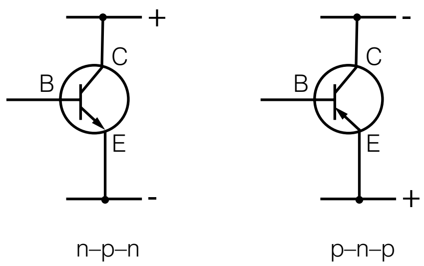


Figure 5.45 *Transistor biasing.*

The arrow on the emitter of a transistor tells you the direction of the conventional current it will allow through. Considering the emitter–collector as the through route, this means that the polarity of the battery across the device must be as in Figure 5.45.

Despite this there should be no current flowing from C because the first junction, the one from collector to base, is reverse biased.

While this is true, it turns out that a small current via the base terminal somehow seems to 'unblock' it, and in those circumstances current will flow from C to E.

Current sent into the base has only one possible route out. The junction towards the collector is reverse biased, so it all has to escape via the emitter (Figure 5.46).

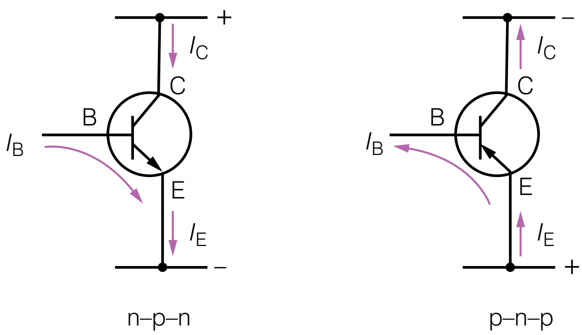


Figure 5.46

By Kirchhoff's current law, (the total current entering any junction in a circuit equals the current leaving it) $I_E = I_C + I_B$.

In terms of voltages, the base has to be at a voltage intermediate between that of the emitter and the collector. Provided V_{BE} is greater than about 0.6 V, the base–collector junction will conduct, so a base current will flow.

Transistor characteristics

We will consider just the n–p–n transistor because that is the more common one. By its characteristics we mean how one variable affects a second one, all other variables staying constant.

In most applications, the signal at the base is used to control what is received at the collector. Two leads are needed to feed an input into a transistor, however, and two leads will be needed for the output as well. A transistor has only three terminals, so one of them will inevitably be common to both. It can be any of the three, so the circuitry can be classified as common emitter, common base or common collector (Figure 5.47).

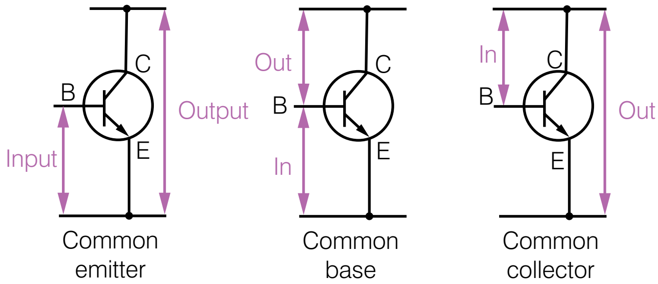


Figure 5.47 *Transistor circuitry.*

These diagrmas do not depict actual circuits – although the input and output are essentially currents, a resistor in the lead between the positive line at the top and the collector will enable the changing voltage across it to be used as an output. Such a common-emitter arrangement is shown in Figure 5.48.

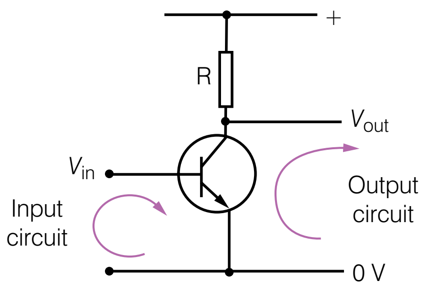


Figure 5.48 *A common emitter.*

Demonstration: Transistor as amplifier

To show how the transistor can be used as an amplifier, connect a microphone to the base of a transistor and a speaker in the collector–emitter part of a circuit (Figure 5.49).

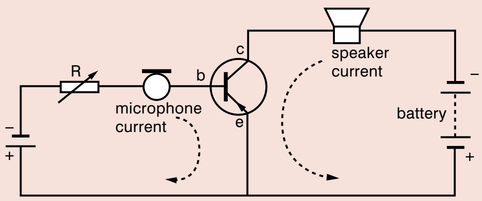


Figure 5.49 *Amplification.*

Such a circuit will cause the very small signal produced by the microphone to become loud enough to be heard through the speaker. In other words, the transistor amplifies the changes in the base current. A very small change in the base current (of the order of μA) causes a much larger change in the collector current (mA).

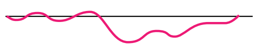


Figure 5.50a *Input to base (microphone).*

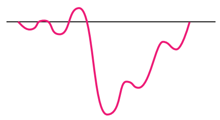


Figure 5.50b *Collector current (speaker).*

If a CRO trace of the signals in the two circuits can be made, you will be able to observe that the variations in the speaker signal are similar in shape but larger than those of the microphone. Figure 5.50a gives an example of variations of the input current (microphone) while Figure 5.50b shows the variations in the output current (speaker). You can see that the signal has been amplified and an estimate of the current gain can be made.

$$\text{Current gain } (h_{fe}) = \frac{\text{transistor output (collector current)}}{\text{transistor input (base current)}} = \frac{I_C}{I_B}$$

Variation in collector current (I_C) with increasing voltage (V_{CE})

One important characteristic is how the collector current (I_C) varies as an increasing voltage V_{CE} is placed across the transistor from emitter to collector. The base current I_B must be held constant, but the graph is usually plotted for more than one base current (Figure 5.51).

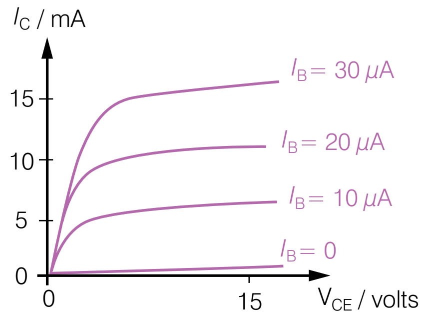


Figure 5.51 *Current through semiconductor plotted against voltage.*

You can see from this that, so long as there is a sufficient voltage between the collector and the emitter, the dominant influence on the collector current is what is being fed into the base. Notice that the base current is small – with the transistor represented by the graph, a change of 10 μA in the base current would produce a change of around 5 mA in the collector current.

Response of collector current to changes in base current

With a transistor which amplifies well, the graph will be nearly a straight line.

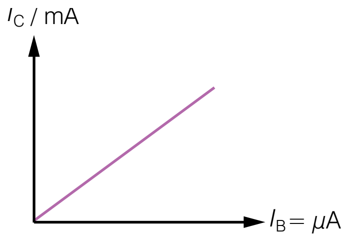


Figure 5.52 *Collector current against base current.*

Transistor as switch and as amplifier

The circuit of Figure 5.53 can be used to demonstrate the behaviour of a transistor. A potential divider is used to raise or lower the voltage at the base to control the base current. If there is no base current, the milliammeter records no collector current either. Move the slider on the potential divider up and down – the base current rises and falls, and the movement on the milliammeter seems to track whatever the base current is doing but with much larger variations.

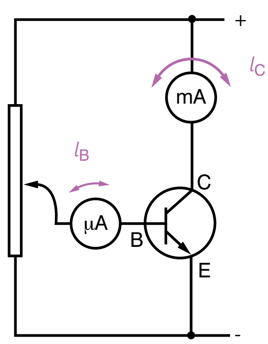


Figure 5.53 *Transistor circuit.*

This demonstration also illustrates two of the important functions of a transistor – it can act as an amplifier, and it can serve as a switch. All transistors can do both jobs, but they are usually designed for one purpose or the other. For amplifiers a linear response is all-important, whereas for a switch what matters is that it does it very quickly.

To act as a switch a small current into the base will switch on a larger current at the collector. This is very much what a **relay** might also do. A transistor may not be able to switch on and off really large currents, but otherwise it has the advantages of no moving parts (and hence reliability) and rapid action.

A transistor might be able to switch off in a time of around 10^{-8} s, while a relay may require something nearer a second. You may feel a second is not too long to wait, but to perform a task in a calculation may call for over a million switching operations, one after another. A second for each is too long then!

KEY WORDS

relay *an electrically-operated switch*

A simple amplifier

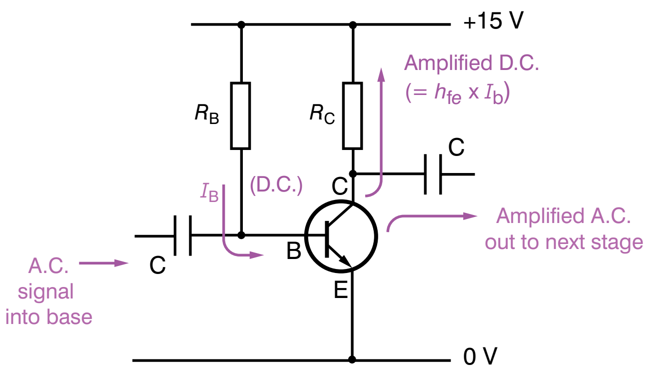


Figure 5.54 *A simple amplifier.*

In Figure 5.54, notice first that the resistor R_B supplies the quiescent base current (that is, a steady d.c. before any input signal is applied). The base–emitter junction is a forward-biased diode, so we may assume that when it is conducting, the voltage across it will be about 0.6 V. If we wish for a base current of, say, 10 µA we can calculate from $R = \frac{V}{I}$ what base resistor is needed. A voltage across it which will be 15–0.6 V has to send 1×10^{-5} A through it – check that R_B will have to be 1.44 MΩ.

Then notice how the input signal is fed in via a capacitor. None of the direct current from R_B can go that way, but the a.c. signal can pass through – and then faced with a choice of the route through to the emitter of the transistor (a low-resistance forward biased diode) or through the 1.44 MΩ, virtually all the signal enters the base.

Now look at the resistor R_C. The collector has to be joined to the positive line to provide a return route for the quiescent current. The value of R_C must be big enough but not too big. It has to be sufficiently large to cause most of the amplified a.c. signal to go instead through the capacitor which links it to the next stage – the larger the better in that respect. The d.c. component of the collector current has to pass through R_C, however, and for that to happen there must be a voltage drop given by $I_C R_C$ across it. The top of R_C is at the +15 V of the supply, so V_{CE} across the transistor must fall by that amount – and we know that if R_C is too big this will make V_{CE} so low that the transistor will stop behaving in a linear fashion or even cease to conduct altogether.

Figure 5.55 is the graph of how I_C varies with V_{CE} for the transistor considered in Figure 5.54. We can see that over its working range a change of 10 μA in the base current will cause a change of around 5 mA in the collector current. The gain (h_{fe}) of the transistor is roughly linear and about 5×10^{-3} A/10×10^{-6} A, which works out to be 500.

We have already chosen a quiescent base current of 10 μA, so the amplitude of the a.c. signal which is to be added must not exceed that value (if it did, the base current would go negative at one of the peaks which would cause the transistor to turn off at that time thereby distorting the output). We are therefore expecting total base currents which will never exceed 20 μA at the most.

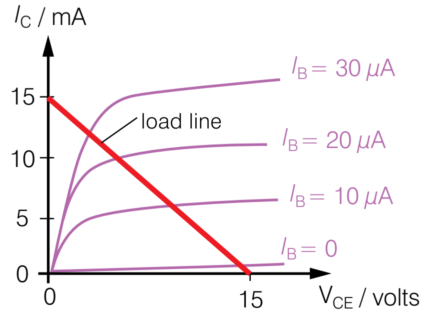


Figure 5.55

The graph now has what we call a load line added. This line shows what the actual value of V_{CE} across the transistor will be – for each value of I_C it is the 15 V of the supply less $I_C R_C$ at that moment. Only when $I_C = 0$ does the full 15 V appear across the transistor.

The straight line must therefore start at (15 V, 0 mA) and is positioned by eye. What is required is that it cuts the I_B = 20 μA line still on an acceptably linear part of its curve. It does not matter that when I_B = 30 μA the voltage has dropped so low that the characteristic is off its linear working range, because we do not intend I_B to get as large as that.

Using the load line drawn, therefore, let us see what is the largest workable value for R_C. Looking at the other intercept, at 15 mA the voltage across the transistor would drop to zero (though it would of course turn off a little before then). A resistor with 15 mA through it and 15 V across it gives a value for R_C of $V/I = 15/15 \times 10^{-3}$ = 1000 Ω. Any value greater than that would mean that in order to handle the expected collector current, the voltage at the collector would drop too low so the output would distort or switch off.

A simple amplifier with negative feedback

Feedback means taking some of the output signal and feeding it back to the input. Negative feedback means that as the output signal rises, a portion is fed back in such a way as to make the input go down a bit and therefore cause the output to drop (Figure 5.56).

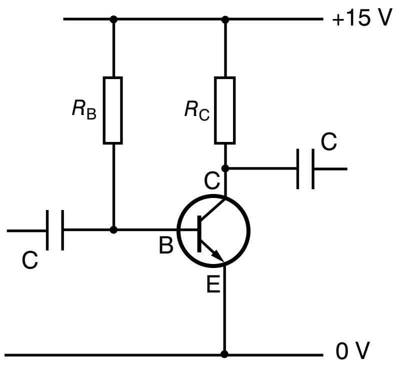


Figure 5.56 *A simple amplifier with negative feedback*

At first sight this might appear undesirable, since it is bound to limit the amplification produced. It does, however, provide useful stability to the transistor. Semiconductors conduct better as they warm up, but this poses a risk: if they conduct better they might pass a larger current, which causes them to heat up further, which makes them conduct more. A good circuit design guards against what could lead to destruction of the semiconductor.

The easiest way to do this is to add a resistor R_E in the emitter line.

To understand how this feedback resistor R_E does its job, suppose the transistor is suffering a surge in the output collector current I_C. Consider the following sequence of events.

1. The current I_E into the emitter must rise too (since $I_E = I_B + I_C$ and the base current is very small).
2. The p.d. across R_E must rise according to $V = IR$.
3. The voltage at emitter E must therefore rise further above 0 V.
4. When the transistor is conducting, like any forward-biased diode the voltage at B will be about 0.6 V higher than that at E, so as the voltage at E rises so must that at B.
5. A higher voltage at B means a smaller p.d. across R_B, so the base current falls.
6. Therefore the collector current I_C drops.

A voltage-operated switch

This circuit is essentially that of an amplifier, but with one difference – the input voltage is either zero or it is a pre-set value (6 V, perhaps), but never anything in between. This means that the transistor is either switched off or it is turned on (Figure 5.57).

The purpose of R_B is to limit the base current to a suitable level, since when 6 V is at the input the only other resistance present is the forward-biased junction.

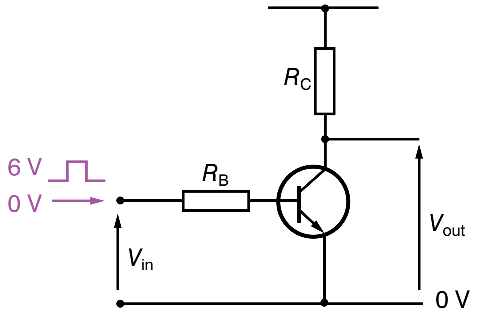


Figure 5.57 *A voltage-operated switch.*

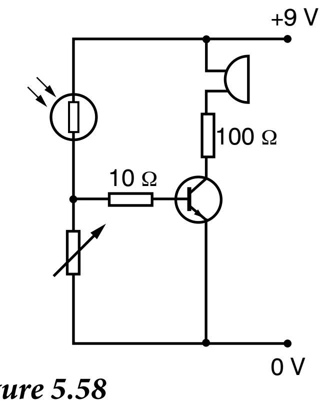


Figure 5.58

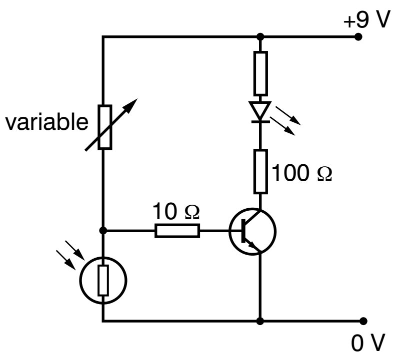


Figure 5.59

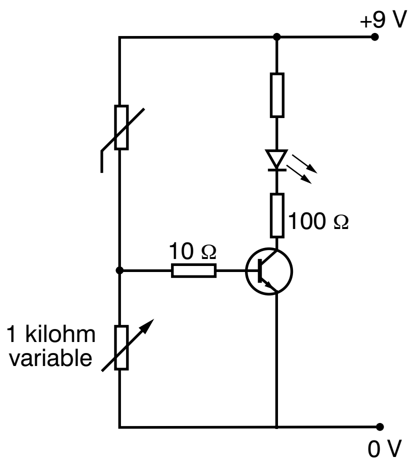


Figure 5.60

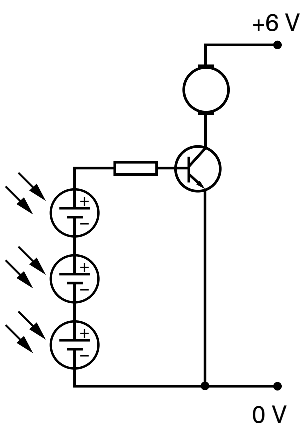


Figure 5.61

This idea is employed in logic circuits, which are discussed later. We say that the input is either a logic '0' or a logic '1', according to whether the voltage is either absent or present.

Using LDR, LED, thermistor, photovoltaic cell and transistor

LDR, LED and transistor

The LDR has a very high resistance when light levels are low, preventing current flowing in the circuit in which it is connected. When light levels rise, the resistance of the LDR drop, allowing current to flow to the base of the transistor.

In the circuit shown in Figure 5.58, the transistor allows current to flow in the circuit containing the buzzer when the LDR is illuminated. This circuit could be the basis for a burglar alarm – if a light shone on the sensor when a room was supposed to be dark, an alarm would sound.

In the circuit shown in Figure 5.59, the transistor allows current to flow in the circuit containing the LED when the LDR receives no light. This circuit could, for example, switch lights on as it gets dark in the evening.

Thermistor, LED and transistor

The thermistor has a very high resistance at low temperatures so, in the circuit shown in Figure 5.60, no current flows to the base of the transistor at low temperatures, and no current is able to flow in the circuit containing the LED. The resistance of the thermistor decreases as the temperature increases allowing current to flow to the base of the transistor. This allows current to flow through the LED.

This circuit could give a visual (or audible if a buzzer were used in place of the LED) warning of overheating.

Photovoltaic cells and transistor

If no light falls on the photovoltaic cells in Figure 5.61, no current is generated in that circuit, and so no current flows to the base of the transistor and no current is therefore able to flow in the circuit containing the motor.

When light shines on the photovoltaic cell, a current flows to the base of the transistor and allows current to flow in the circuit in which a motor is connected.

The motor in such a circuit could be used to operate a blind to cover a window if the sun shone onto the sensor.

Logic gates

Logic gates are tiny silicon chips on which are etched combinations of transistors and resistors. They typically have two inputs and one output. What happens at the output is determined by the situation at those inputs. An example is the AND gate, which we shall consider first.

The AND gate

The power supply needed will be a steady smooth d.c. of around 5 V. Each gate must be connected to this power supply by two wires, which are not normally indicated on the diagrams: one at +5 V, and the other at 0 V (Figure 5.62).

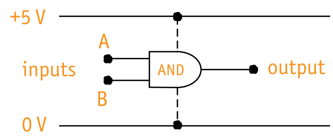


Figure 5.62 *AND gate – the dotted lines are not usually shown.*

Each input may either be joined to the 0 V line, in which case we say the input is a '0', or it may be joined to the +5 V line, in which case the input is a '1'. In Figure 5.63 input A is a '1' and B is a '0'.

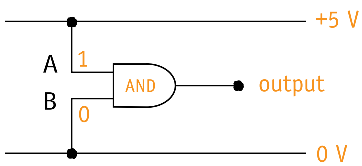


Figure 5.63 *Input A is '1' (one) and input B is '0' (zero).*

The output may be a '0' (so it is at the same voltage as the 0 V line) or it may be a '1' (and so acts as if it was the +5 V line). The output of a logic gate is shown by its truth table which is a list of all possible input combinations, showing what you get each time at the output.

Figure 5.64 shows the truth table for the AND gate. As its name suggests, the only way to get a '1' at the output is to have a '1' at input A and a '1' at input B. In Figure 5.63, you can see that input A is '1', input B is '0'. From the truth table you can see that the output for these input values is '0'.

INPUT A	INPUT B	OUTPUT
0	0	0
0	1	0
1	0	0
1	1	1

Figure 5.64 *Truth table for AND gate.*

KEY WORDS

logic gate *an electronic device that performs a logical operation on two inputs and produces a single logic output*

You can think of an AND gate as a bit like two switches in series (Figure 5.65). Both switches must be placed in the '1' position to make the output 'live' (that is, at +5 V).

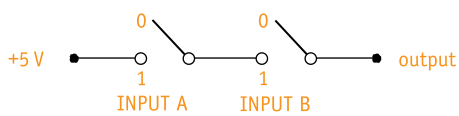


Figure 5.65 *You can think of an AND gate as two switches in series.*

As the inputs are voltages and not currents, all you have to do is touch the input contact onto the 'live' (+5 V) line for a '1', or touch it on the earth line for a '0'. The input resistance is huge, so in effect no current flows into the gate.

The OR gate

The OR gate may not be quite what you would expect from its name. If the AND gate could be pictured as two switches in series, the OR gate behaves like the same two switches in parallel (Figure 5.66b).

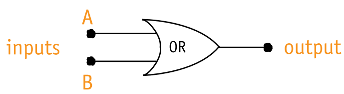


Figure 5.66a *The symbol for an OR gate.*

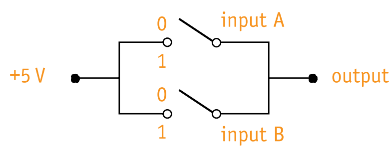


Figure 5.66b *You can think of an OR gate as two switches in parallel.*

The truth table for the OR gate is shown in Figure 5.67.

INPUT A	INPUT B	OUTPUT
0	0	0
0	1	1
1	0	1
1	1	1

***Figure 5.67** Truth table for OR gate.*

You might not have expected the last line in the table from what we normally understand by the word 'or', but if you think of those two switches in parallel (Figure 5.66b), you should soon be able to work out the correct truth table. The output will always be at +5 V except when both switches are at 'zero'.

The NOT gate

Yet another gate is the NOT gate. This has a single input, and its output is always the opposite. A '0' at the input means a '1' at the output, and vice versa. Figure 5.68 shows the symbol for a NOT gate.

If you are puzzled as to how a '1' can come out when nothing is going in, you are forgetting that the gate also has connections to the power supply which are not shown. If the input is connected to earth (0), therefore, the +5 V of the 'live' line appears at the output. The truth table for the NOT gate is shown in Figure 5.69.

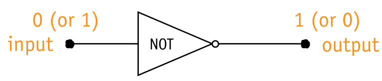


***Figure 5.68** The symbol for a NOT gate.*

INPUT	OUTPUT
0	1
1	0

***Figure 5.69** Truth table for NOT gate.*

The NAND gate and the NOR gate

There are two final gates to consider – the NAND gate and the NOR gate. These are just the AND gate and the OR gate respectively, but with the output inverted. Instead of a '0' there is a '1', and instead of a '1' a '0'. Their symbols are shown in Figure 5.70 and the little circle at the output indicates that each value is reversed.

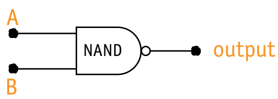


***Figure 5.70a** Symbol for a NAND gate.*

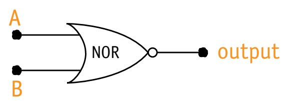


***Figure 5.70b** Symbol for a NOR gate.*

The four truth tables are summarised in Figure 5.71.

			Output		
INPUT A	**INPUT B**	**AND**	**NAND**	**OR**	**NOR**
0	0	0	1	0	1
0	1	0	1	1	0
1	0	0	1	1	0
1	1	1	0	1	0

***Figure 5.71** Truth table for NAND and NOR gates.*

Nobody would expect you to sit and learn that table by heart. You should be able to work out how the AND and OR gates behave, and then obtain the other two simply by reversing their outputs.

Combinations of logic gates

More than one logic gate may be combined to increase the range of control tasks that can be performed.

As an example, consider the arrangement shown in Figure 5.72. How does it behave?

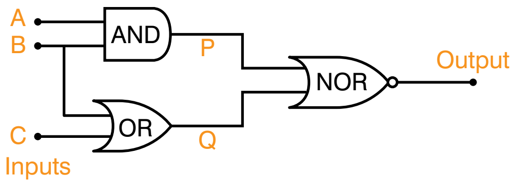


Figure 5.72 *Combining different logic gates.*

Start by preparing a table with columns for all possible combinations of inputs for the corresponding logic states at intermediate points P and Q, and finally for the output.

There are three inputs this time, which will give $2^3 = 8$ different combinations of 0s and 1s. If you know your binary notation, the neatest way is to write them down is in sequence by listing in turn the binary equivalents of 0 to 7. Figure 5.73 shows the truth table filled in with these figures.

INPUT A	INPUT B	INPUT C	POINT P	POINT Q	OUTPUT
0	0	0			
0	0	1			
0	1	0			
0	1	1			
1	0	0			
1	0	1			
1	1	0			
1	1	1			

Figure 5.73 *Initial truth table for logic gate combination shown in Figure 5.72.*

Once the inputs are entered, work out the states of points P (from inputs A and B) and Q (from inputs B and C) and record them in the table (Figure 5.73).

Then use P and Q as inputs to the final gate in order to find the output.

Check that in this case the result is as shown in Figure 5.74.

INPUT A	INPUT B	INPUT C	POINT P	POINT Q	OUTPUT
0	0	0	0	0	1
0	0	1	0	1	0
0	1	0	0	1	0
0	1	1	0	1	0
1	0	0	0	0	1
1	0	1	0	1	0
1	1	0	1	1	0
1	1	1	1	1	0

Figure 5.74 *Final truth table for logic gate combination shown in Figure 5.72.*

Worked example 5.4

Consider the arrangement shown in Figure 5.75. How does it behave?

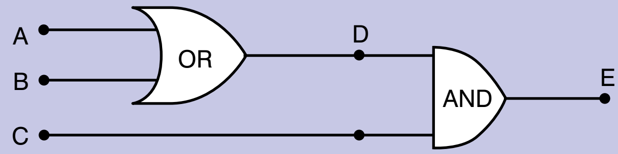


Figure 5.75

Start by preparing a table with columns for all possible combinations of inputs (A, B and C). There are three inputs this time, which will give $2^3 = 8$ different combinations of 0s and 1s. As seen above, the neatest way is to write them down is in sequence by listing in turn the binary equivalents of 0 to 7. Figure 5.76 shows the table filled in like this.

INPUT A	INPUT B	INPUT C	POINT D	POINT E
0	0	0		
0	0	1		
0	1	0		
0	1	1		
1	0	0		
1	0	1		
1	1	1		

Figure 5.76 *Intermediate truth table for logic gate combination shown in Figure 5.75.*

Now fill in the column for the logic state at intermediate point D (remember to use columns for Input A and Input B to work out the results for point D), and finally for the output, E (using the columns for Input C and point D). Figure 5.77 shows the completed table.

INPUT A	INPUT B	INPUT C	POINT D	POINT E
0	0	0	0	0
0	0	1	0	0
0	1	0	1	0
0	1	1	1	1
1	0	0	1	0
1	0	1	1	1
1	1	1	1	1

Figure 5.77 *Final truth table for logic gate combination shown in Figure 5.75.*

The action of logic gates

In this section we will give some circuits involving logic gates and say what they are meant to do. Your task each time is to try to explain how they work.

A simple burglar alarm

In the circuit drawn in Figure 5.78, two doors and two windows are equipped with micro-switches such that if they are opened the switch closes and the signal from it changes from a 0 to a 1.

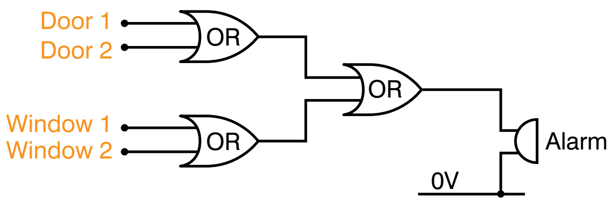


Figure 5.78 *A circuit for a simple burglar alarm.*

The switches are connected to one of a pair of two-input OR-gates, so if any of the inputs becomes a 1 the alarm will sound.

A thermostat for a hot water tank

In the circuit drawn in Figure 5.79,

(a) two contacts are fixed near the top of the tank and when submerged (i.e. when the water is deep enough) the water completes a circuit between the contacts, signalling that the water is deep enough,

(b) a thermistor (set to the required temperature) is immersed in the water in the tank; it signals if the temperature of the water is too cold.

The water heater will only be switched on if both (a) the tank is full of water AND (b) the water is too cold.

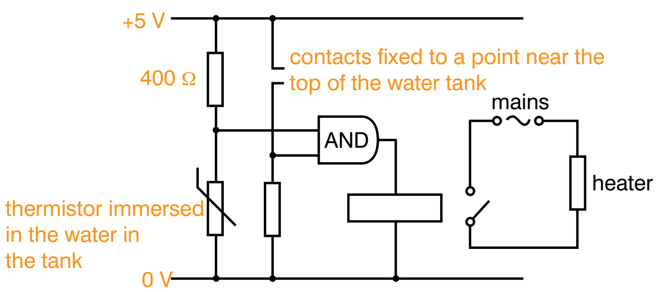


Figure 5.79 *A thermostat for a hot water tank.*

What the output of logic gates can do – using a relay

When there is an '0' at the output, it acts as if it was the 0 V line. When there is a '1', it behaves like part of the +5 V line and is capable (in principle, at least) of lighting a bulb (Figure 5.80).

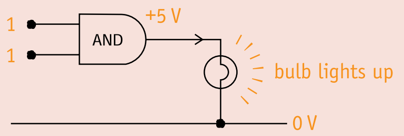


Figure 5.80 *When the output is a '1' a current can flow.*

The only snag is that the tiny silicon chip cannot really handle the sort of current needed to make the filament of a bulb white-hot, but it can cope with the much smaller current needed to make a light emitting diode glow.

There is one other thing that the small current which a gate will supply is capable of doing – it can operate a relay. The output of the gate feeds current to an electromagnet. An entirely separate circuit contains a switch made of iron. The electromagnet will attract the iron switch towards it against the pull of a spring, and this action either turns the separate circuit on or switches it off (Figure 5.81).

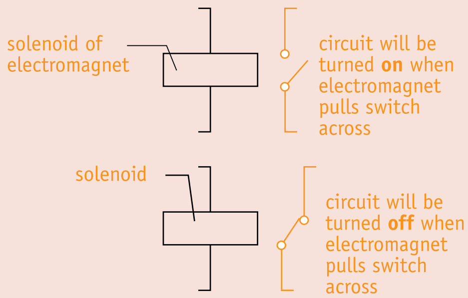


Figure 5.81 *How a relay works.*

The circuit symbol for each type of relay is shown in Figure 5.81. Figure 5.82 shows an AND gate whose inputs are both '1', so it will turn on a powerful electric motor.

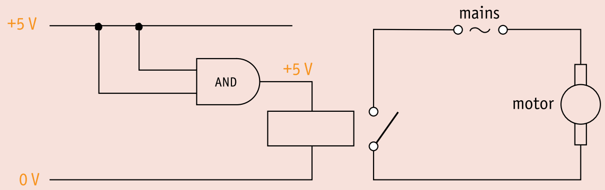


Figure 5.82 *How an AND gate can operate a relay to switch a motor on and off.*

A security lock to get 'behind the counter' at a bank

In the circuit drawn in Figure 5.83, button A is on the bank manager's desk and button B is by the door of the bank, on the outside.

The solenoid which pulls the bolt in, unlocking the door, is only activated if both button A AND button B are pressed and held down.

The warning buzzer sounds if button B is pressed AND button A is NOT pressed.

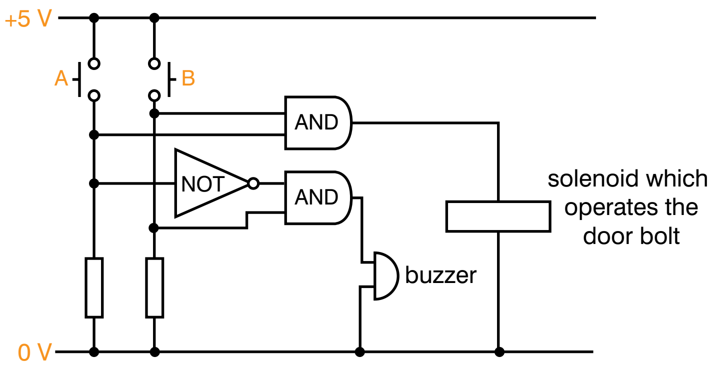


Figure 5.83 *A security lock circuit.*

An automatic plant waterer

In the circuit drawn in Figure 5.84:

(a) Contacts sensing nmoisture content are placed in the soil. These generate a signal when the soil is damp. As the system is required to switch the pump on when the soil is dry, the signal passes through a NOT gate.

(b) A light sensor, consisting of a photovoltaic cell, generates a voltage when light falls on it, but as the system is required to switch on when it is dark, the signal passes through a NOT gate.

(c) Switch S is a manual override. Pressing this switch will switch on the pump whatever the conditions.

The solenoid switching on the pump to water the plants will come on if the soil is dry AND it is dark OR if the manual override switch is pressed.

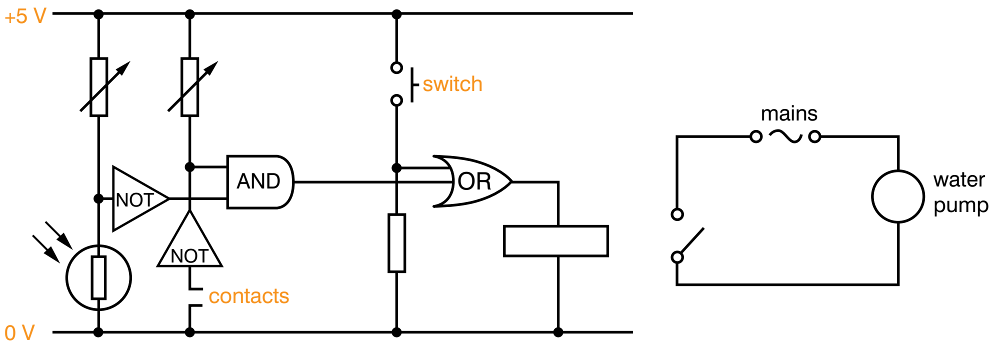


Figure 5.84 *An automatic plant watering system.*

Integrated circuits

Logic gates are grouped together in circuits – such as the gate shown in Figure 5.85a.The tiny silicon chip is encased in resin and is connected to the outside world via its 14 metal 'legs' (Figure 5.85b). Note that the gates are identified just by their symbols. It is a 'quad 2-input NAND' meaning that it contains four separate NAND-gates, each with an input A and an input B. A single +5 V connection and a single 0 V one serve all four gates inside.

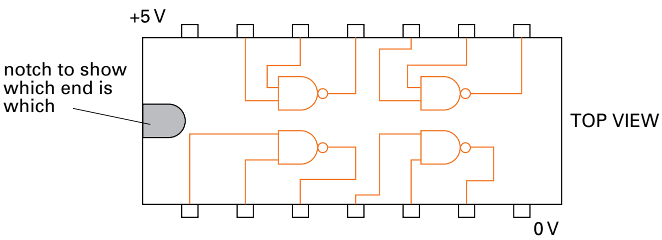


Figure 5.85a *Layout.*

Figure 5.85b *Photo of a quad 2-input NAND.*

Summary

- With an n–p–n transistor, there is a low input resistance at the base, and the base current flows out of the emitter via the forward-biased junction.
- The gain h_{fe} of the transistor is defined to be $\frac{\Delta I_C}{\Delta I_B}$.
- The collector current is almost as big as the emitter current: $I_E = I_C + I_B$.
- A transistor can be used as an amplifier or as a switch.
- Logic gates work in terms of 1s (a voltage is present) and 0s (the voltage is not present).
- The truth table for a logic gate is simply a list of all the possible input readings with their corresponding output.
- Logic gates can be combined to perform all sorts of tasks relating to control and alarms.

Review questions

1. A transistor has a gain h_{fe} of 200. If a current of 3.0 μA is sent into the base, what size collector current would you expect?
2. Compile your own notes on types of logic gate.
   a) Start with the NOT gate. Give its symbol (just its outline shape this time, without writing the word 'NOT' on it) and its truth table, which for this gate will have only two columns and two lines to it.
   b) Now do the same for the AND gate. When you come to the truth table, see if you can get all four lines right before you check – if it helps, add a note comparing it to two switches in series (Figure 5.65).
   c) Do the NAND gate next. Again, working from the AND gate, try to write down the truth table before looking.
   d) Next do the OR-gate in the same way. This time, if it helps, add a note comparing it to two switches in parallel (Figure 5.66b).
   e) Finally repeat for the NOR gate.
3. a) In Figure 5.86, does the output of the logic gate need to be a '0' or a '1' in order for the tiny bulb to light? Explain.
   b) What inputs will be needed at A and B to achieve this?
4. Complete the truth table for the arrangement of gates shown in Figure 5.87.

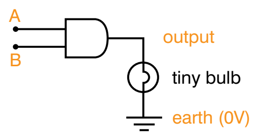


Figure 5.86

A	B	C	D	E
0	0	1		
1	0	0		
0	1	1		
1	1	0		

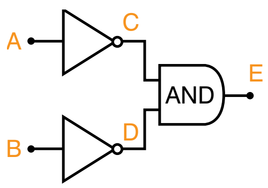


Figure 5.87

End of unit questions

1. A battery is connected to the y terminals of an oscilloscope and the spot deflects 3 cm vertically. The sensitivity of the oscilloscope is set at 1 V/cm. What is the p.d. across the battery?
2. Boron has only three electrons in its outer shell. Explain why doping a silicon crystal with a small amount of boron makes it a p-type semiconductor.
3. What is the difference between d.c. obtained from a battery and that obtained from an alternating supply when you add a diode in series?
4. What do you understand by n-type and p-type semiconductors?
5. The input of a logic gate may be either a '0' or a '1'. Explain what that means.
6. a) Name the type of logic gate shown in a Figure 5.88a. Write down its truth table.

   b) Show that when its two inputs are joined together as in Figure 5.88b it will act as a NOT gate.

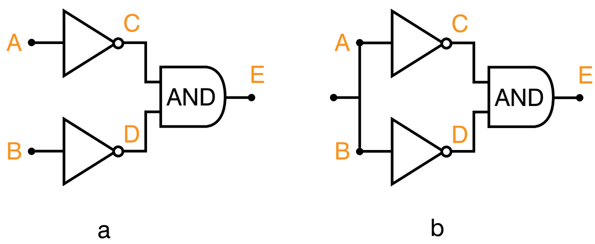


Figure 5.88

Electromagnetic waves and geometrical optics

Unit 6

Contents

Section	Learning competencies
6.1 Electromagnetic waves (page 195)	• Explain how electromagnetic waves are produced. • Describe the nature of electromagnetic waves. • Compare mechanical and electromagnetic waves. • Draw diagrams to represent transverse waves. • Use straight lines to represent the direction of energy flow (rays). • Identify that electromagnetic waves emitted by the Sun have a wide continuous range of frequencies (and wavelengths). • Explain some uses of electromagnetic radiation.
6.2 Reflection of light (page 199)	• Explain what is meant by the rectilinear propagation of light. • State the laws of reflection. • Perform experiments to test the laws of reflection using a plane mirror. • Use the laws of reflection to explain how images are formed in a plane mirror. • Find the position of a virtual image produced by a plane mirror using a ray tracing method. • Use the laws of reflection to solve problems. • Give examples of the uses of plane mirrors. • Distinguish between concave and convex mirrors. • Identify the meanings of: principal axis, principal focus, radius of curvature, magnification in relation to concave and convex mirrors. • Distinguish between real and virtual images. • Apply the appropriate sign convention when using mirror equations. • Find the position and nature of the image formed by a concave and a convex mirror using the mirror equation and a ray tracing method. • Use the relation magnification $= \frac{S_i}{S_o} = \frac{h_i}{h_o}$ to solve problems. • Give examples of the uses of curved (concave and convex) mirrors.
6.3 Refraction of light (page 211)	• Define the term refraction. • State the conditions in which refraction occurs. • Define the refractive index of a material. • Use Snell's law to solve simple problems. • Use the formula refractive index $= \frac{\text{real depth}}{\text{apparent depth}}$ to find the refractive index of a liquid and a solid in the form of a rectangular glass block. • Perform experiments to test the laws of refraction. • Draw a diagram representing the passage of light rays through a rectangular glass block. • Give examples of observations that indicate that light can be refracted.

Contents

Section	Learning competencies
	• Identify that the passage of a ray of light through a parallel-sided transparent medium results in the lateral displacement of a ray. • Define the critical angle θ_c. • Explain, with the aid of a diagram, what is meant by critical angle and total internal reflection. • Identify the conditions necessary for total internal reflection to occur. • Perform calculations involving critical angle and total internal reflection. • Describe how total internal reflection is used in optical fibres. • Distinguish between convex and concave lenses. • Identify the meaning of: principal focus, principal axis, focal point, radius of curvature, magnification in relation to converging and diverging lenses. • Apply the appropriate sign convention when using thin lens equations. • Find the position and nature of the image formed by a convex and concave lens using the thin lens formula and a ray tracing method. • Define the power of a lens. • Explain how the image is formed due to combination of thin lenses. • Draw a ray diagram to show how images are formed by lenses in a simple microscope and a simple telescope. • Compare and contrast the structure and functions of the human eye and the camera. • Describe how the human eye forms an image on the retina for different object distances. • Identify some defects of the eye and their correction with lenses. • Explain what is meant by the dispersion of white light to produce a spectrum. • Identify that the passage of a ray of light through a triangular transparent prism results in a deviation of a ray.

6.1 Electromagnetic waves

By the end of this section you should be able to:

- Explain how electromagnetic waves are produced.
- Describe the nature of electromagnetic waves.
- Compare mechanical and electromagnetic waves.
- Draw diagrams to represent transverse waves.
- Use straight lines to represent the direction of energy flow (rays).
- Identify that electromagnetic waves emitted by the Sun have a wide continuous range of frequencies (and wavelengths).
- Explain some uses of electromagnetic radiation.

KEY WORDS

transverse waves *waves that oscillate perpendicular to the axis along which they travel*

mechanical waves *travel through a material or substance, with time. They have frequency, period, wavelength and amplitude*

electromagnetic waves *transverse waves produced when a magnetic field and an electric field are at right angles to each other*

Types of wave

There are two types of wave: transverse and longitudinal. In this unit we shall be considering **transverse waves**.

All waves are produced by vibrations. In a transverse wave, the vibrations are at right angles to the direction of movement, as shown in Figure 6.1.

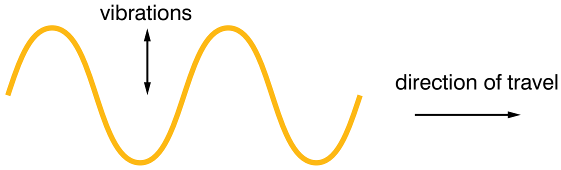


Figure 6.1 *Transverse waves.*

Mechanical waves travel through some medium (material), with time. They have a frequency, period, wavelength and amplitude. Mechanical waves transport energy and not material.

Electromagnetic waves

The vibrations in **electromagnetic waves** come from electric and magnetic fields. Electromagnetic waves are produced when a magnetic field and an electric field are at right angles to each other. Charges that are accelerated in an electric and magnetic field produce electromagnetic waves. They carry energy and momentum that may be transferred to matter with which they interact.

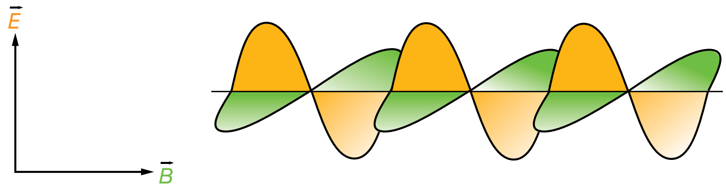


Figure 6.2 *Electromagnetic waves.*

Electromagnetic waves are transverse waves. Electromagnetic waves do not need a medium through which to travel – they can travel through a vacuum.

Activity 6.1: Modelling an electromagnetic wave

You will need to stand with a group of 20 of your fellow students in a line. Number yourselves from 1 to 20. Those of you who have an odd number are to represent the magnetic field. Those of you who have an even number are to represent the electric field.

Those of you who are representing the magnetic field should hold your hands out in front of you, while those of you who are representing the electric field should hold your hands in the air, as shown in Figure 6.3.

You learnt about Faraday's law in Unit 4. Remember that if you increase an electric field, you will induce a magnetic field and that decreasing the electric field will induce a magnetic field in the opposite direction.

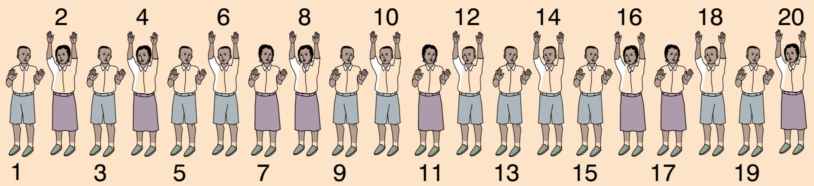


Figure 6.3 *Students representing an electromagnetic wave.*

Demonstrate Faraday's law using your hands: you can represent changes in the electric field by changing the height of your hands in the air and represent changes in the magnetic field by turning through 180°.

Activity 6.2: Compare mechanical and electromagnetic waves

In a small group, draw up a table to compare mechanical and electromagnetic waves. Make sure you understand what it is that is moving in each case.

KEY WORDS

amplitude *the maximum distance a wave moves above or below the base line*

frequency *the number of complete waves passing a given point in a second*

wavelength *the distance between successive peaks or troughs on a wave*

rays *straight lines extending from a point*

speed *distance travelled per unit time*

Representing transverse waves

You can draw diagrams to represent transverse waves such as the one in Figure 6.4.

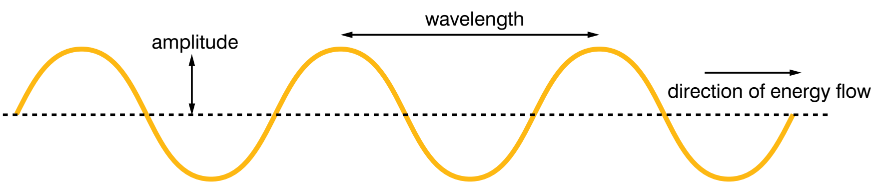


Figure 6.4 *Transverse wave diagram.*

The **amplitude** is the distance that the wave moves above or below the base line. The **frequency** of a wave is the number of complete waves passing a given point in a second. The **wavelength** is the distance between successive equivalent points (usually taken as peaks or troughs) on a wave. You can use straight lines (**rays**) to represent the direction of energy flow along the wave.

Relationship between frequency, wavelength and speed

The frequency, wavelength and **speed** of a wave are related by the formula:

speed = frequency × wavelength

In symbols this is written as:

$v = f\lambda$

The units need to be:

- speed: m/s
- frequency: Hertz
- wavelength: metres.

DID YOU KNOW?

Heinrich Hertz (1857–1894) was a German physicist who was the first to demonstrate satisfactorily the existence of electromagnetic waves by building an apparatus to produce and detect VHF or UHF radio waves.

The unit of frequency for waves is named after him.

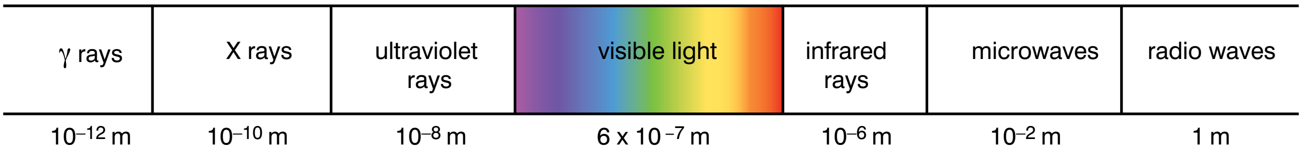


Figure 6.5 *Electromagnetic spectrum.*

Worked example 6.1

A water wave has a wavelength of 2 cm and a frequency of 15 Hz. Find its speed.

v (m/s)	λ (m)	f (Hz)
?	2 cm	15

First we need to make sure that the given values are in the correct units.

Wavelength = 2 cm = 0.02 m

Frequency = 15 Hz

Substitute these values into the equation:

speed = frequency × wavelength

gives

speed = 15 × 0.02
= 0.3 m/s

Activity 6.3: Uses of electromagnetic radiation

In a small group, design a poster showing uses of electromagnetic radiation. You may also like to include some hazards of electro-magnetic radiation – do some research to find out about gamma rays, for example.

KEY WORDS

electromagnetic spectrum *the entire frequency range of electromagnetic waves*

Electromagnetic waves emitted by the sun

The Sun emits electromagnetic waves with a wide range of frequencies and wavelengths. This is generally referred to as the **electromagnetic spectrum**, as shown in Figure 6.5.

In a vacuum, all electromagnetic waves travel at a speed of 3×10^8 m/s. They have different properties (and therefore uses) because, as shown in Figure 6.5, they have different wavelengths (and therefore frequencies).

Some uses of electromagnetic radiation

Visible light is the most important part of the electromagnetic spectrum to everyday life, even though its wavelengths are only a small part of the spectrum.

X-rays are used to take pictures of inside the body to show any bone fractures. They are absorbed more by bone (which is denser than the surrounding muscles).

Infrared radiation is used in infrared cameras, which are used in rescue operations to detect the presence of bodies.

Microwaves and radio waves are used for communications – for example, radio and telephone signals.

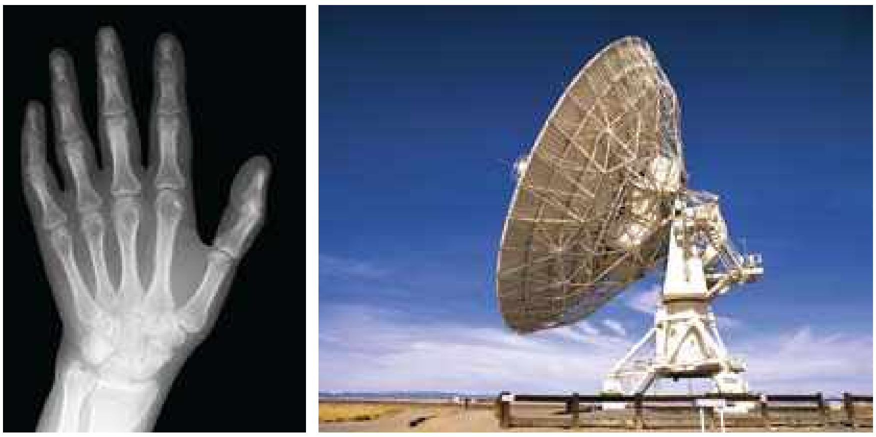


Figure 6.6 *Some uses of electromagnetic radiation.*

Summary

- Electromagnetic waves are produced when electric fields and magnetic fields are at right angles to each other.
- Electromagnetic waves are transverse waves that can travel in a vacuum.

- Mechanical waves transfer energy and need a medium through which to travel. Electromagnetic waves can travel through a vacuum. They have energy and momentum that may be transferred to bodies with which they interact.
- You can draw diagrams to represent transverse waves and use straight lines to represent the direction of energy flow (rays).
- Electromagnetic waves emitted by the Sun have a wide continuous range of frequencies (and wavelengths) – this is called the electromagnetic spectrum.
- X-rays are used to take pictures of inside the body to show any bone fractures. They are absorbed more by bone (which is denser than the surrounding muscles).
- Infrared radiation is used in infrared cameras which are used in rescue operations to detect the presence of bodies.
- Microwaves and radio waves are used for communications – for example, radio and telephone signals.

Review questions

1. Explain how electromagnetic waves are produced.
2. Describe the nature of electromagnetic waves.
3. Compare mechanical and electromagnetic waves.
4. Draw a diagram to represent a transverse wave. Show its amplitude, frequency and wavelength. Use straight lines to represent the direction of energy flow (ray).
5. What is the relationship between speed, frequency and wavelength for a transverse wave?
6. Give some types of electromagnetic waves on the electromagnetic spectrum.
7. What is the frequency of an X-ray with a wavelength of 1×10^{-10} m?

6.2 Reflection of light

By the end of this section you should be able to:

- Explain what is meant by the rectilinear propagation of light.
- State the laws of reflection.
- Perform experiments to test the laws of reflection using a plane mirror.
- Use the laws of reflection to explain how images are formed in a plane mirror.
- Find the position of a virtual image produced by a plane mirror using a ray tracing method.
- Use the laws of reflection to solve problems.
- Give examples of the uses of plane mirrors.
- Distinguish between concave and convex mirrors.
- Identify the meanings of: principal axis, principal focus, radius of curvature, magnification in relation to concave and convex mirrors.

- Distinguish between real and virtual images.
- Apply the appropriate sign convention when using mirror equations.
- Find the position and nature of the image formed by a concave and a convex mirror using the mirror equation and a ray tracing method.
- Use the relation magnification $= \frac{s_i}{s_o} = \frac{h_i}{h_o}$ to solve problems.
- Give examples of the uses of curved (concave and convex) mirrors.

Rectilinear propagation of light

In the previous section, you learnt that visible light is a form of electromagnetic radiation, or electromagnetic wave. **Rectilinear propagation** of light simply means that light waves travel in straight lines.

KEY WORDS

rectilinear propagation *a wave property by which waves travel in straight lines*

Activity 6.4: Investigating the reflection of light

Set up the apparatus as shown in Figure 6.7.

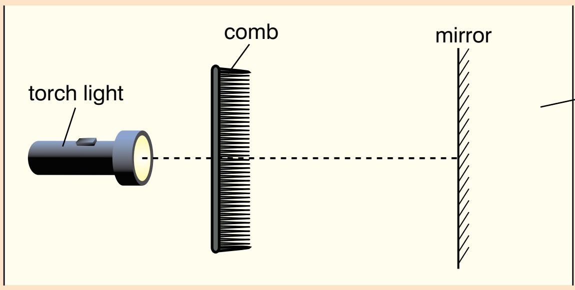


Figure 6.7 *Apparatus for investigating the reflection of light.*

Break every other tooth of the comb. Draw the rays that appear to be coming from the mirror on the paper.

Activity 6.5: Natural examples of reflection

In a small group, make a list of where you see reflection in nature.

The laws of reflection

Before we go on to state the laws of reflection, we need to sort out a few points.

The first is that angles are not measured from the mirror surface itself, but from a construction line called the normal. This line is at right angles to the mirror surface and should be drawn dotted. Thus a beam of light hitting the mirror at an angle of 30° would look as shown in Figure 6.8.

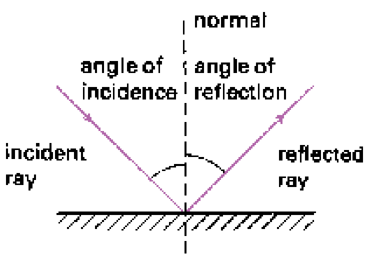


Figure 6.8 *Angles are measured from the normal.*

Remember that **angles are always measured from the normal**. It is easy to forget this and measure the wrong angle.

The beam of light on its way to the mirror is called the incident ray. It hits the mirror ('is incident upon the mirror') at a particular angle of incidence. After that the reflected ray leaves the mirror at the appropriate angle of reflection. Figure 6.9 illustrates this.

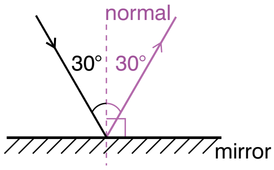


Figure 6.9 *Reflection at a mirror.*

Now we can consider how light behaves when it is reflected. The results of Activity 6.4 can be summed up in two laws. The first law of reflection is straightforward:

The angle of reflection is equal to the angle of incidence.

In other words, if light hits a mirror at 30° to the normal, it will leave the mirror at 30° to the normal too.

At first sight this one law says everything. It takes a bit of thought to realise that this is not so. There is actually an endless number of possible rays of light at 30° to the normal, and they form a kind of cone round the normal, as in Figure 6.10.

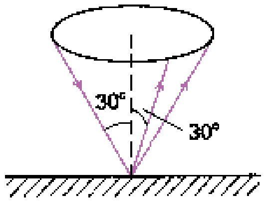


Figure 6.10 *The reflected ray is the one opposite to the incident ray.*

The ray we mean is the one opposite to the incident ray. This is expressed more scientifically in the second law of reflection:

The reflected ray lies in the plane which contains the incident ray and the normal.

The word 'plane' means flat surface, like a sheet of paper. An ordinary flat mirror is more correctly known as a **plane mirror**.

Activity 6.6: Making a simple periscope

Figure 6.11 shows the principles of a simple periscope.

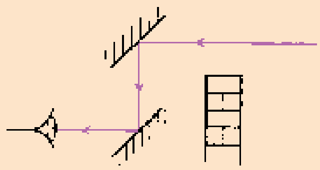


Figure 6.11 *Principles of a simple periscope.*

If you can acquire two small plane mirrors, you could make one for yourself. Each mirror needs to be inclined at an angle of 45°, so you will need to design a cardboard tube to support them.

Why you cannot see your reflection in a sheet of paper

A sheet of white paper is a good reflector, yet it is useless if you wish to see to comb your hair. Why?

We say it is a good reflector because of all the light energy landing on it, perhaps 90% or more is reflected off again. It does not form a visible **image**, however, because that light gets scattered in all directions.

The key to this behaviour lies in the surface of the paper. At a microscopic level the paper is all humps and bumps whereas a mirror is smooth. The consequence of this is shown in Figure 6.12.

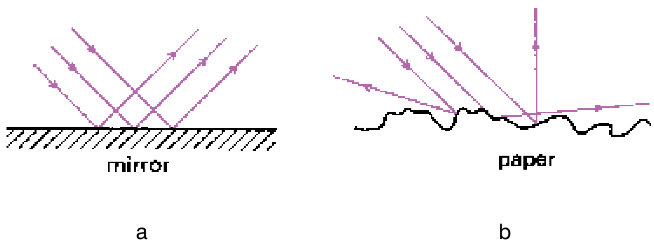


Figure 6.12 *a) Regular and b) diffuse reflection.*

Notice that with the paper the laws of reflection apply as much as with a mirror. The light still gets reflected off at the same angle that it hit the surface: it is just that the surface points in different directions.

A shiny black surface is just the opposite of the paper. It is smooth, but it absorbs much of the light that lands on it. The small amount of light that gets reflected leaves in a regular manner and so you may be able to see a faint reflected image in it.

KEY WORDS

plane mirror *a mirror whose surface lies in a plane*

image *an optically formed reproduction of an object*

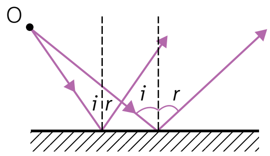


Figure 6.13

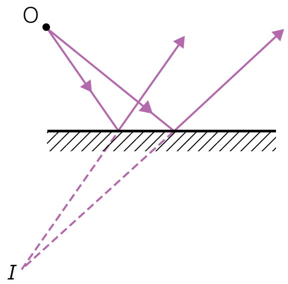


Figure 6.14

Using the laws of reflection to explain how images are formed in a plane mirror and using a ray tracing method to find the position of the image

Consider the mirror shown in Figure 6.13, with the object at point O.

The first law tells us that the angle of incidence is equal to the angle of reflection, so in the diagram $i = r$.

Now trace the reflected rays beyond the mirror using dotted lines, as shown in Figure 6.14.

The point where the dotted lines meet is the position of the image, I.

Worked example 6.2

The diagram shows a small source of light 'O'. Marked on the diagram are the edges of a cone of light which eventually enters the eye of a person placed roughly as shown in Figure 6.15.

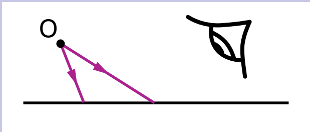


Figure 6.15

a) Using a sharp pencil and a ruler, draw a diagram which looks similar but larger. Leave some space below the mirror. Don't put the eye in the diagram at first.

b) With the aid of a protractor, accurately draw in the normals (dotted lines) where the edges of the cone of light meet the mirror. For each one, carefully measure the exact angle of incidence (from the normal!) and accurately draw the paths of the light after reflection.

c) The light entering the eye is diverging (i.e. spreading out). To locate the spot from which the light seems to be coming, extend each of your two reflected lines backwards to where they meet (again use dotted lines for this: no light actually takes that route).

a)

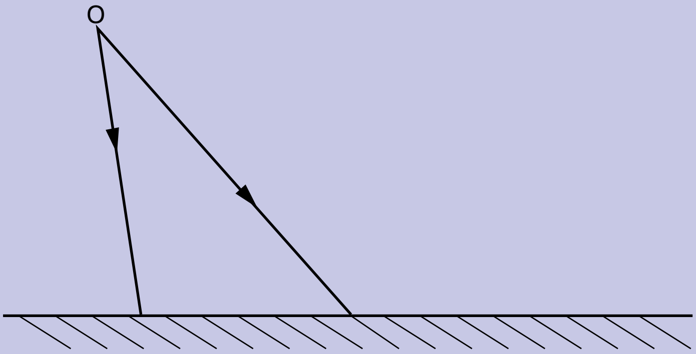


Figure 6.16

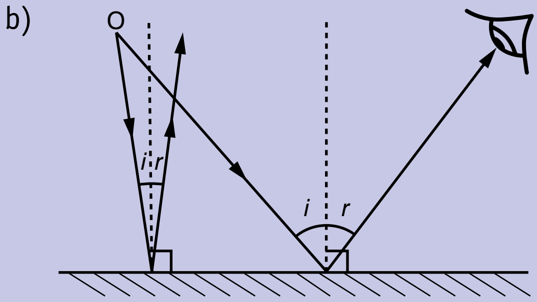


Figure 6.17

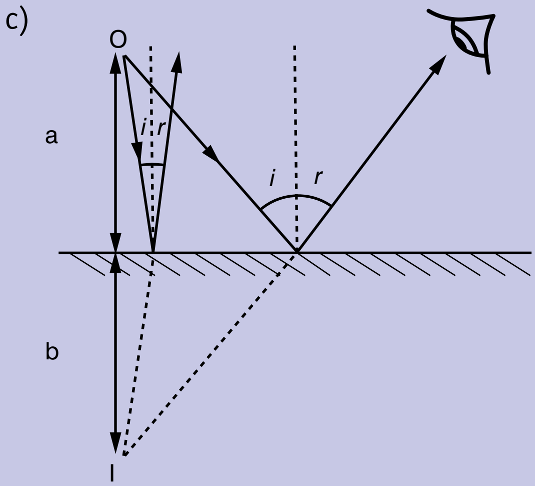


Figure 6.18

Activity 6.7: The uses of plane mirrors

Work in a small group to make a list of all the uses of plane mirrors that you can think of.

Concave mirrors

In the worked example above you found that, in a plane mirror, the image is erect, virtual, laterally inverted, and the same size as the object.

We will now consider curved mirrors. You have probably seen reflections in polished metal spheres. A shiny tablespoon forms reflections from its inside surface and from its outside surface. All these reflections are images, things we see even though the object is somewhere else. This section considers how these images are formed.

First, we must sort out a few names. All the mirrors in this section are spherical mirrors, which means that they are shaped like part of the surface of a sphere; their curvature in all directions is the same.

We will start with a **concave mirror**. Its shape and the way it behaves are shown in Figure 6.19, and you should soon be able to convince yourself that this is exactly in line with the basic laws of reflection.

KEY WORDS

concave mirror *a mirror with a reflecting surface that bulges inwards, away from the light source*

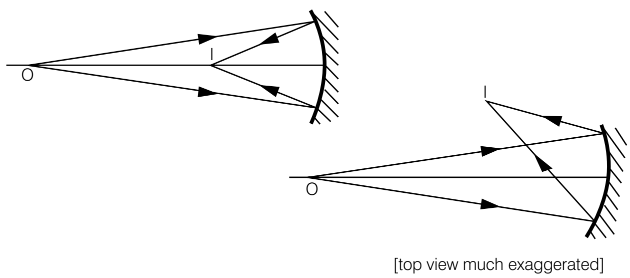


Figure 6.19 *Light reflecting off a concave mirror.*

Activity 6.8: Investigating the behaviour of a concave mirror

What happens as the object comes closer? The mirror may then still use the laws of reflection to bring the light to a focus, but if so it will no longer be at the focus (that is, the **principal focus**).

Experiment and find out. Figure 6.20 shows a possible experimental set-up.

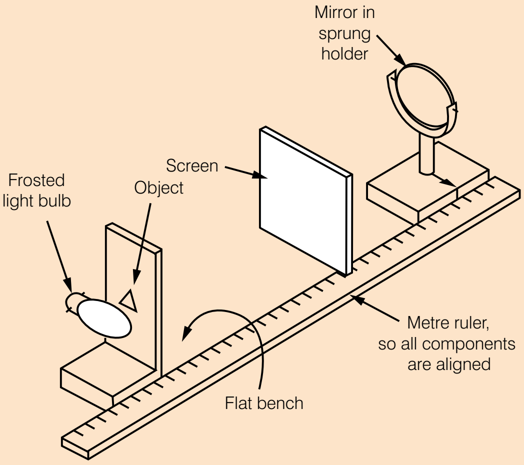


Figure 6.20 *Possible experimental set-up.*

The object is a triangular-shaped hole in a wooden board, and fine detail in the object is provided by a wire gauze across the hole. This fine detail helps in judging when the image is in sharp focus on the screen. (An alternative is a transparent plastic ruler with clear markings, held in a clamp.) To ensure that the image is bright enough to see, the object is illuminated from the back, just like the slide in a slide projector.

Start with the object as far away from the mirror as you can, around 1 m perhaps. The task now is to find where a screen must be to obtain a clear focused image on it. Immediately a difficulty crops up: the screen will have to be put right in the way of the light on its way to the mirror (as shown in Figure 6.19).

The easiest way round the difficulty is to have the mirror at a slight angle – as shown in the lower diagram of Figure 6.19, which is much exaggerated.

Once the apparatus is set up, test it fully. You should find you can get a sharp upside-down image on the screen. As the object gets closer to the mirror the screen has to be moved further and further back, while the image gets larger and larger.

KEY WORDS

focus *the point at which light rays converge*

convex mirror *a mirror with a reflecting surface that bulges outwards, towards the light source*

focal length *the distance from the centre of a curved mirror to the principal focus*

Concave mirrors and magnification of images

In Activity 6.8, you found that as the object approached the principal focus of the mirror, the screen had to be moved a great distance away. Once you are inside that focus there is nowhere for the screen to go to receive an image on it – the mirror no longer forms a real image anywhere.

Using your own face as the object, you will find that when you are close enough the mirror can still be used as a looking glass. It forms a virtual image, just as a plane mirror did. Conveniently, this image is the right way up and there is one big difference compared to the plane mirror: the image is larger than the object. It has been magnified. The field of view is less than that in a plane mirror. The magnification produced, however, means that a converging mirror is sometimes used as a men's shaving mirror.

Convex mirrors

Convex mirrors reflect the light so it spreads out. If the incoming light is a parallel beam, we take as the focal length the distance from the principal focus '*f*' to the centre of the mirror (see Figure 6.21).

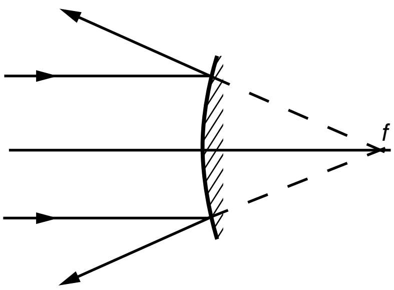


Figure 6.21 *Light reflecting off a convex mirror.*

A convex mirror cannot form a real image, so you will never need a screen with it. It can act as a looking glass, however, and you will be able to use it to see a virtual image, which seems to be a short way behind it.

This image has the following properties:

1. It is always the right way up.
2. It always has a **magnification** of less than 1. Although everything appears smaller, this means that the complete field of view is a wide one.
3. It has a striking clarity. This is because normally our eyes can only focus clearly on a limited range of distances – we can look at a fly crawling up a window pane or we can look at the view outside but not both at the same time. In the mirror all objects from nearby to the far distance are seen as images compressed into the short distance between the mirror and its focus. Now our eyes can view them all clearly at the same moment – a unique experience.

Terms used in concave and convex mirrors

Figure 6.22 summarises the terms you have already met in relation to concave and convex mirrors. It shows the **principal axis**, **principal focus** and **radius of curvature** of a concave and a convex mirror. The magnification is given by $\frac{\text{height image}}{\text{height object}}$.

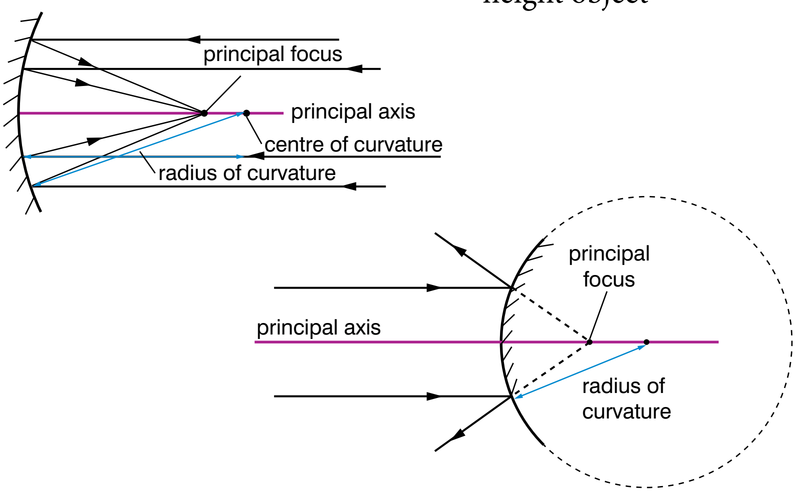


Figure 6.22 *Terms used in relation to concave and convex mirrors.*

Distinguish between real and virtual images

We have already met the terms real and virtual images. Now we shall explain what these terms mean.

When you can focus the light from an image on a screen, the image is called a **real image**. If no screen happens to be there, the light actually passes through that point, and an alternative way to view it is simply to stand back from it and look.

KEY WORDS

magnification *the ratio between the height of an image and the height of the object*

principal axis *the line passing through the optical vertex and centre of curvature of the face of a curved mirror*

principal focus *the point at which all light reflecting from a curved mirror converges*

radius of curvature *the radius of the sphere that forms the basic curve of a concave mirror*

real image *an image that can be captured on a screen*

KEY WORDS

virtual image *an image that cannot be captured on a screen*

Notice that you must be sufficiently far back from the image to be able to see it. Placing your eye at the image position itself is like trying to read this book by resting your eyeball on the page.

The image produced by a plane mirror is different. We can stand back and view the image just the same, but this time it has been produced by what you could describe as a trick. If you put a screen behind the mirror, of course, no image will be formed on it because the light never actually started from that point. We call this a **virtual image**.

Activity 6.9: The uses of concave and convex mirrors

Some cars use diverging mirrors. Plane mirrors are also available and are sometimes fitted. It is not so easy to judge distances in them, but diverging mirrors provide a far better field of view (see Figure 6.23).

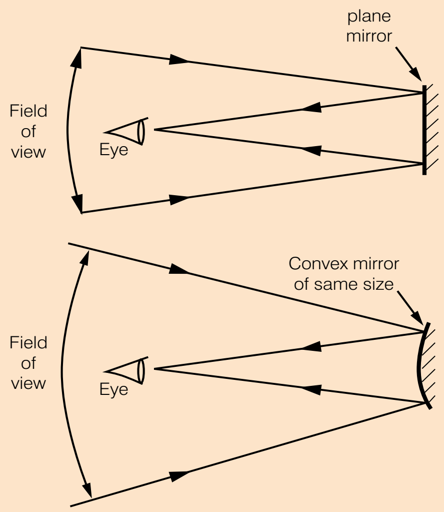


Figure 6.23 *Light reflecting off a car mirror compared to a plane mirror.*

Figure 6.24 *Light reflecting off a bus's mirror.*

In a small group, make a list of as many uses for concave and convex mirrors as you can. Think about their magnification and their field of view if it helps you to think about where they may be used.

Why we think of light from distant objects as parallel

It is not quite mathematically parallel, but as far as measurements are concerned it is parallel enough for the accuracy we are seeking.

The more distant the better, of course. Only a point object at infinity would give perfectly parallel light, but an object several metres away (let alone the Sun at about 1.5×10^{11} m distance) is adequate for our measurement.

Activity 6.10: Determining the focal length of a concave mirror

Imagine a test beam of parallel light being shone along the principal axis of a concave mirror. It will be brought to a focus (the *principal focus, F*) of the mirror. The distance from the principal focus to the vertex of the mirror is its *focal length, f*. The greater the curvature of the mirror, the closer will be its principal focus and the shorter its focal length will be (see Figure 6.25).

The easiest way to find the focal length is to shine parallel light onto the mirror and see where this light is brought to a focus. Hold the mirror in a stand and take a screen – a sheet of paper will do if need be. Light coming from distant objects is sufficiently parallel for this purpose, so move the paper towards the mirror until an image of a window at the far side of the room (or, even better, a tree in the distance) is formed clearly on the screen. It will be upside down, quite small and – to many people's surprise – in colour.

Measure the distance from the screen to the vertex of the mirror, and you have its focal length.

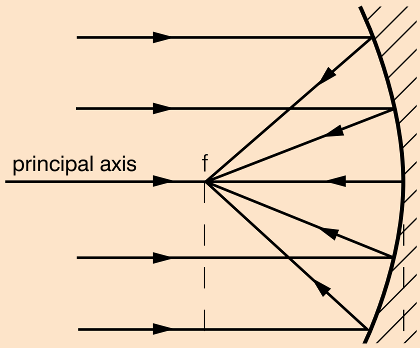


***Figure 6.25** Light coming to a focus after reflection from a convex mirror.*

The mirror equation

If you know the focal length of a converging mirror, it is possible to do a calculation to give answers to questions such as 'If I place an object at a given distance from the mirror, whereabouts will I have to put the screen, and how big will the image be?'

The notation used is:

f = the focal length of the mirror.

s_o = the distance from the object to the centre of the mirror.

s_i = the distance from the centre of the mirror to where the image is formed.

If the image is virtual, then we use a negative sign for the distances.

The connection between them is:

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f}$$

Worked example 6.3

A mirror has a focal length of 200 mm (0.20 m). If an object is placed 0.60 m from the mirror, where will the image be formed?

Use $\frac{1}{f} = \frac{1}{s_o} + \frac{1}{s_i}$

f (m)	s_o (m)	s_i (m)
0.20	0.60	?

Putting in the values,

$$\frac{1}{0.60} + \frac{1}{s_i} = \frac{1}{0.20}$$

This gives $1.67 + \frac{1}{s_i} = 5.00$

Solving, $\frac{1}{s_i} = 5.00 - 1.67 = 3.33$

This is $\frac{1}{s_i}$, remember, so to find s_i we must 'turn the answer upside down'.

$s_i = \frac{1}{3.33} = 0.30$ m (30 cm, 300 mm).

Therefore a screen would have to be placed 300 mm back from the mirror.

Finding the position and nature of the image formed by a concave and a convex mirror using the mirror equation and a ray tracing method

Figure 6.26 shows how you can find the position and nature of the image formed by a concave mirror using a ray tracing method.

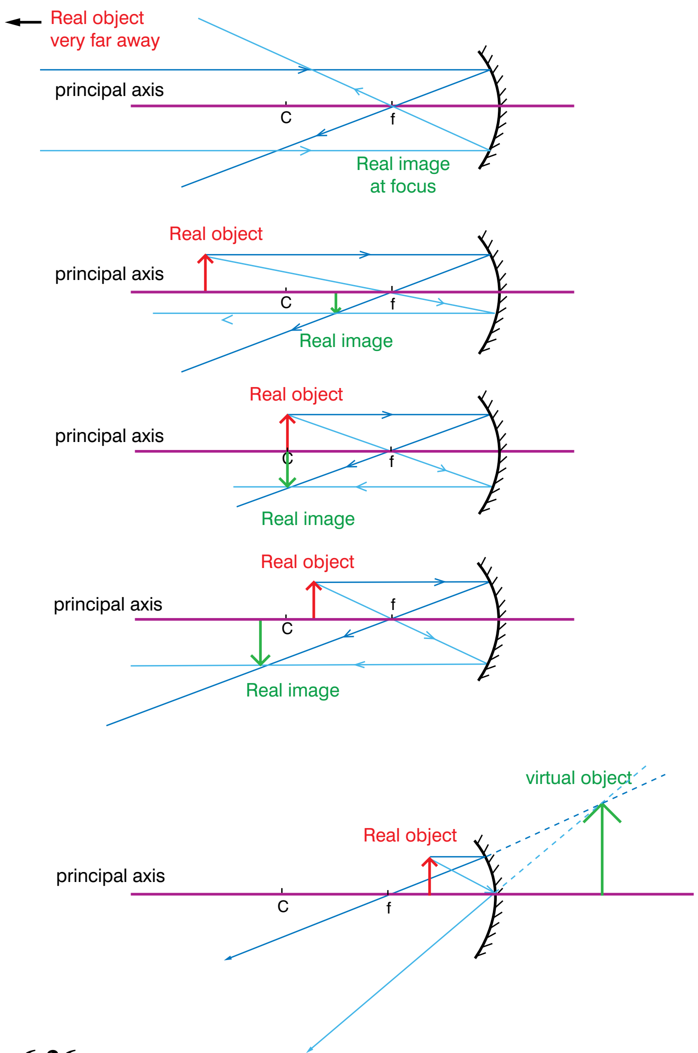


Figure 6.26

Worked example 6.4

You stand 15 cm away from a converging mirror of focal length 20 cm.

a) Work out the distance to the image.

b) Calculate the magnification of the image. (For this purpose ignore the minus sign.)

s_o (cm)	s_i (cm)	f_i (cm)
15	?	10

a) Use $\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f}$

$$\frac{1}{s_i} = \frac{1}{f} - \frac{1}{s_o} = \frac{1}{20} - \frac{1}{15}$$

$$= -\frac{1}{60}$$

$s_i = -60$ cm

b) Use magnification $= \frac{s_o}{s_i} = \frac{60}{15} = 4$

The magnification relationship for mirrors

Magnification is defined to be:

$$\frac{\textbf{the height of the image}}{\textbf{the height of the object}}$$

It is just a number, and will have no units. We can calculate what it will be like this:

Magnification $= \dfrac{h_i}{h_o}$

Thus if $h_o = 0.60$ m and $h_i = 0.30$ m, the magnification can be expected to be:

$$\frac{h_i}{h_o} = \frac{0.30}{0.60} = 0.50.$$

Notice that the magnification here turns out to be less than 1. The image will be only half as tall as the original object – if we used a ruler for the object, the millimetre divisions in the image would be only 0.5 mm apart. We still refer to this as magnification, however.

This table summarises what you have learnt about curved mirrors.

Quantity	Concave mirror	Convex mirror
Focal length f	+ve sign	–ve sign
Object distance s_o	In front of mirror +ve sign	Behind mirror –ve sign (virtual)
Image distance s_i	In front of mirror +ve sign (real)	Behind mirror (virtual)
Magnification m	Image upright +ve sign	Image inverted –ve sign

Summary

- Rectilinear propagation of light simply means that light travels in straight lines.
- The laws of reflection are:

  The angle of reflection is equal to the angle of incidence.

  The reflected ray lies in the plane which contains the incident ray and the normal.
- An example of a use of plane mirrors is seeing your reflection as you do your hair.
- Concave and convex mirrors curve as shown in Figure 6.27.

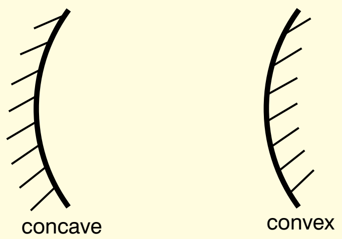


Figure 6.27

- Figure 6.28 shows the principal axis, principal focus, radius of curvature for concave and convex mirrors.

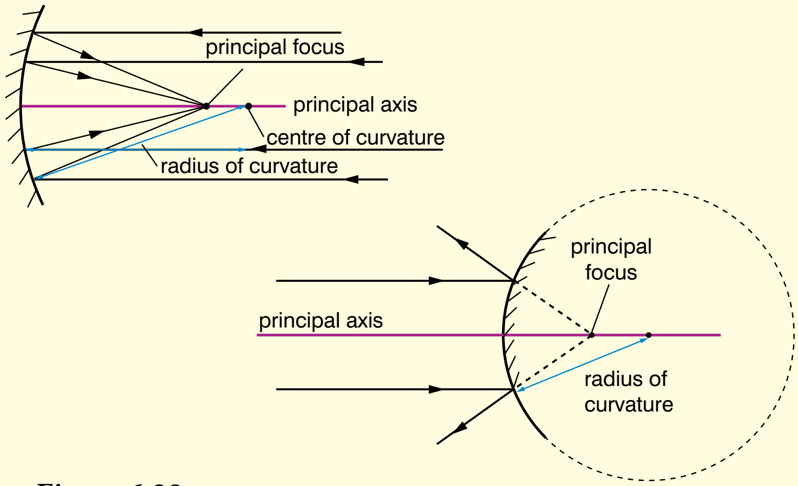


Figure 6.28

- Real images can be focused on a screen. Virtual images cannot be focused on a screen.
- When using the mirror equation, virtual images have a negative sign.
- You can find the position and nature of the image formed by a concave and a convex mirror using the mirror equation and a ray tracing method.
- Magnification = $\frac{s_i}{s_o} = \frac{h_i}{h_o}$
- Examples of the uses of curved (concave and convex) mirrors are shaving mirrors and wing mirrors on vehicles.

Review questions

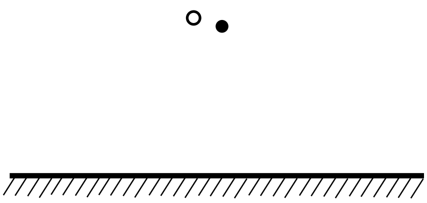


Figure 6.29

1. Explain what is meant by the rectilinear propagation of light.
2. State the laws of reflection.
3. Use the laws of reflection to explain how images are formed in a plane mirror.
4. Find the position of the image produced by the plane mirror in Figure 6.29 using a ray tracing method.
5. An object is 10 cm from a plane mirror. The angle of incidence on a plane mirror is 35°. Draw a diagram to show the situation and work out:
   a) the angle of reflection
   b) the position of the virtual image.
6. Give examples of the uses of plane mirrors.
7. Distinguish between concave and convex mirrors.

8. Draw a diagram to show the meanings of principal axis, principal focus and radius of curvature in relation to concave and convex mirrors.
9. Define magnification in relation to concave and convex mirrors.
10. Distinguish between real and virtual images.
11. State the mirror equation and the sign convention that you need to apply when using it.
12. Find the position and nature of the image formed by the concave and convex mirrors in Figure 6.30 using:
    a) the mirror equation
    b) a ray tracing method.

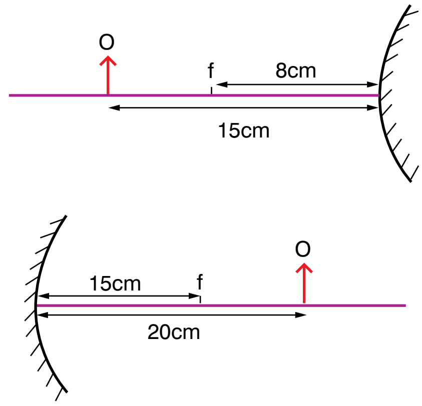


Figure 6.30

6.3 Refraction of light

By the end of this section you should be able to:

- Define the term refraction.
- State the conditions in which refraction occurs.
- Define the refractive index of a material.
- Use Snell's Law to solve simple problems.
- Use the formula refractive index = $\frac{\text{real depth}}{\text{apparent depth}}$ to find the refractive index of a liquid and a solid in the form of a rectangular glass block.
- Perform experiments to test the laws of refraction.
- Draw a diagram representing the passage of light rays through a rectangular glass block.
- Give examples of observations that indicate that light can be refracted.
- Identify that the passage of a ray of light through a parallel-sided transparent medium results in the lateral displacement of a ray.
- Define the critical angle θ_c.
- Explain, with the aid of a diagram, what is meant by critical angle and total internal reflection.
- Identify the conditions necessary for total internal reflection to occur.
- Perform calculations involving critical angle and total internal reflection.
- Describe how total internal reflection is used in optical fibres.

KEY WORDS

refraction *the change in direction of travel of a light beam as the light crosses the boundary between one transparent medium and another*

- Distinguish between convex and concave lenses.
- Identify the meaning of: principal focus, principal axis, focal point, radius of curvature, magnification in relation to converging and diverging lenses.
- Apply the appropriate sign convention when using thin lens equations.
- Find the position and nature of the image formed by a convex and concave lens using the thin lens formula and a ray tracing method.
- Define the power of a lens.
- Explain how the image is formed due to combination of thin lenses.
- Draw a ray diagram to show how images are formed by lenses in a simple microscope and a simple telescope.
- Compare and contrast the structure and functions of the human eye and the camera.
- Describe how the human eye forms an image on the retina for different object distances.
- Identify some defects of the eye and their correction with lenses.
- Explain what is meant by the dispersion of white light to produce a spectrum.
- Identify that the passage of a ray of light through a triangular transparent prism results in a deviation of a ray.

Figure 6.31

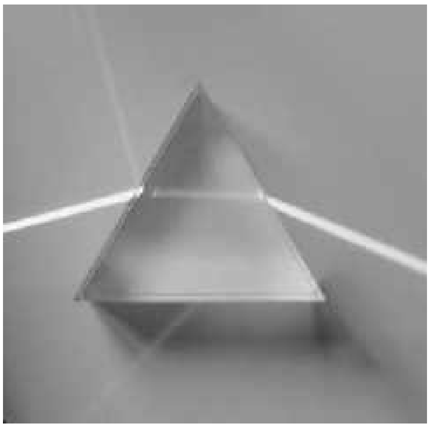


Figure 6.32

Refraction

Something odd is happening in Figure 6.31 and it is obviously being caused by the presence of the water. Reflection cannot account for it, so what is the explanation?

Now look at Figure 6.32. The light is coming from the left, and reaches a triangular glass prism. A little of that light is reflected from each surface of the glass – that is how you may be able to see your reflection in a glass window. Most of the light goes through the glass, however. You can see from the figure that at the boundaries between the air and glass, the light beam undergoes a sudden change in direction. This is what we call **refraction**.

Refraction is the change in the direction of travel of a light beam that occurs as the light crosses the boundary between one transparent medium and another.

This bending of a light beam as it crosses a surface is a very important effect in optics: it is the basic principle behind the working of cameras and telescopes, for example. We must therefore examine refraction in more detail.

Which way does the light get bent?

To work out which way the light gets refracted at a particular surface, it is helpful to draw the normal to the surface (as a dotted line). Having done that, remember that the word 'into' is enough to predict the path of the light. This tells you that as the light goes *in* to the water or the glass, it is bent towards the normal. What this means is shown in Figure 6.33. Notice that the light always crosses the normal.

Light going the other way (coming *out from* the water or the glass) would take the same route in the other direction: as it comes out it is refracted away from the normal.

The only time the light does not get refracted as it crosses the boundary is if it hits the surface 'square on', right along the normal (which will be an angle of incidence of 0°, remember, not 90°).

A few consequences of refraction

We can now begin to understand why we get photographs like those at the start of this section. A careful study of Figure 6.34 provides the answer. The essential thing to realise is this: where something seems to be is determined solely by the direction in which the light is travelling as it *enters* the eye.

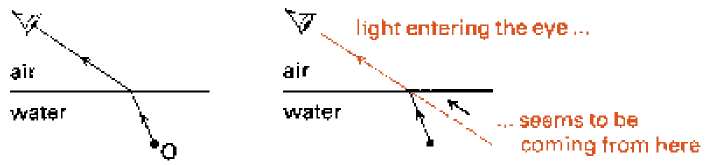


Figure 6.34 *The path of the light as it enters the eye is what determines where the object appears to be.*

If we look up at the night sky, a similar situation means that no star is really where it appears to be unless it happens to be directly overhead. This is due to refraction in the atmosphere. (The boundary does not have to be between two different substances; refraction will occur at a boundary between 'thin' air, as in the upper atmosphere, and the denser air lower down.) Figure 6.35 shows a simple model of the Earth's atmosphere which we can use. At each boundary the light goes *into* denser air and each time it gets refracted towards the normal.

In practice the density of the air increases steadily as you go down, of course. We could improve our model by dividing it into a large number of very thin layers, which would turn the path of the light into a smooth curve. The end result is the same: the star looks higher in the sky than it really is.

Another example of refraction in the atmosphere is the shimmering effect that can be seen in the air above a Bunsen burner or a very hot surface. This is convection in action. As the hot air swirls upwards, its density keeps changing randomly. This causes light that crosses the hot air to be refracted first one way and then the other, so objects constantly appear to be shifting their position slightly.

Activity 6.11: Observing refraction

Take a glass of water. Put a straw or stick into the glass and describe what you see.

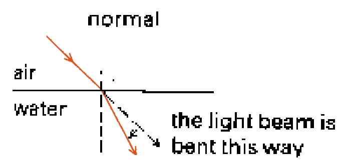


Figure 6.33 *The light always crosses the normal, but is bent towards it.*

Figure 6.35 *The star seems to us to be higher than it really is.*

KEY WORDS

refractive index *a measure of the extent to which a medium refracts light*

Snell's law *whenever light crosses a boundary between two media, the sines of the angles on each side of the boundary bear a constant ratio to each other*

The refractive index and Snell's law

Some materials refract light at their boundary more than others. The extent to which each one does this is measured by its **refractive index**, given the symbol 'n'.

The refractive index is a number larger than 1 such that the greater the number the greater the refraction produced. Water, for example, has a refractive index of 1.33, while that of common glass is just a shade over 1.5. Figure 6.36 illustrates to scale the difference that this makes. The water has the smaller refractive index, so its surface bends the light less.

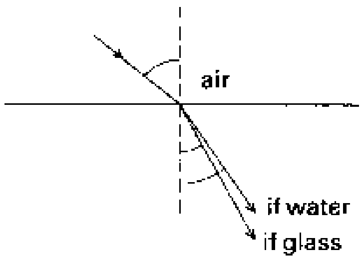


Figure 6.36 The greater the refractive index, the more the light is bent.

What does the number mean? It is a measure of the refraction that occurs at a boundary between the material or 'medium' in question (e.g. glass or water) and a vacuum. (In practice, we use a boundary with air, which behaves virtually the same way as a vacuum.)

Figure 6.37 shows the situation. No arrow has been drawn on the light beam because it follows the same path whichever way it is going. If angle θ_1 is increased to such an extent that its sine is doubled, then angle θ_2 will become larger so that its sine is doubled too.

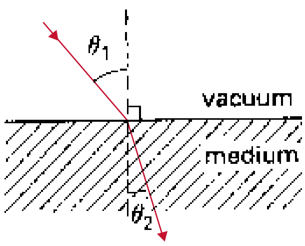


Figure 6.37 Light crossing a boundary.

This relationship was first spotted by a Dutch mathematician called Willebrord Snell back in 1621, and is now known as **Snell's law**:

Whenever light crosses a boundary between two transparent media, the sines of the angles on each side of the boundary bear a constant ratio to each other.

Mathematically, for a particular boundary:

$$\frac{\sin\theta_1}{\sin\theta_2}$$

always comes to the same number regardless of the size of the angles themselves. If that boundary is with a vacuum (or, near enough, with air), this number is the **refractive index** of the medium.

$$n = \frac{\sin\theta_1}{\sin\theta_2}$$

To calculate the path light will take as it crosses a boundary between *any* two media, the simplest way is to use Snell's law in the form:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1}$$

where the notation is as shown in Figure 6.38.

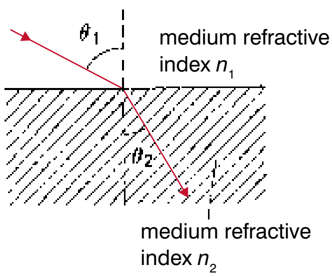


Figure 6.38 The notation for Snell's law.

Since 'n' is more than 1, irrespective of the direction the light is travelling in you always put the sine of the larger angle on top.

Activity 6.12: Discussion

In a small group, discuss these questions.

1. What do you think would happen if glass of refractive index 1.52 is immersed under a liquid whose refractive index is also 1.52?
2. How do you think the speed of light in the glass compares with that in the liquid?

Set up some apparatus to find out!

Why does refraction occur?

Refraction was first noticed with light. Early scientists soon found that the relationship on either side of a boundary was not between the angles themselves, but between their sines. They went on to use their observations to design lenses, and to make microscopes and telescopes.

It has since been discovered that refraction occurs with all types of waves, including ripples on the surface of water and sound waves. The bending is caused by a change in the speed of the wave when it crosses a boundary. If a medium has a high refractive index, that tells us that light entering the medium from a vacuum will undergo a large amount of refraction. It must also mean that the light will be slowed down a lot as it travels from the vacuum into the medium.

The real reason that the refractive index can be defined and measured by the ratio of the sines of the angles is that it is also the ratio of the speeds on either side of the boundary – the faster speed over the slower speed. It can be shown that these two alternative definitions are in fact equivalent, so one follows from the other.

Thus if the speed of light in a vacuum is denoted by c and its (slower) speed in the medium is v, then the refractive index n of the medium will be given by:

$$n = \frac{c}{v}$$

Lateral displacement

Because the two opposite faces of the block are parallel, by geometry the light must meet the second face at the same angle, θ_2, (see Figure 6.40b). This means that as the light leaves the glass it is refracted by the same amount the other way, and so must emerge on a path parallel to its original one as shown in Figure 6.40b. The sideways shift of the beam is called the **lateral displacement**.

The extent of the lateral displacement depends on the angle at which the light is incident on the outer surface of the glass block. At an angle of incidence of 0° (that is, when it hits the block at right angles to its surface) the lateral displacement is of course nil. As the angle of incidence increases, so does the displacement.

Worked example 6.5

Light travelling in air meets the surface of a block of Perspex at an angle of incidence of 50°. It enters the block at an angle of refraction of 31°. Calculate the refractive index of Perspex.

Start by drawing the diagram in Figure 6.39. Mark all the angles to the normal.

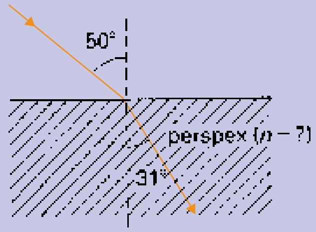


Figure 6.39

$$n = \frac{\sin \theta_1}{\sin \theta_2}$$

$$n = \frac{\sin 50°}{\sin 31°}$$

$n = 1.49$

KEY WORDS

lateral displacement *the perpendicular distance between the pathway of the incident light ray and the one that emerges after refraction from two surfaces of a medium*

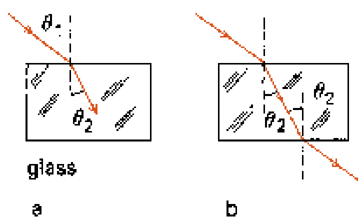


Figure 6.40 *Light entering and leaving a rectangular glass block.*

Activity 6.13: Testing the laws of refraction

Take a piece of plain paper. Set up the apparatus shown in Figure 6.40 so that a narrow beam of light is incident on one face of a rectangular glass block and is refracted as it crosses the boundary.

Trace round the outline of the block. Use a sharp pencil to trace the incident beam of light and the beam of light as it exits the block. Take the paper from under the block and join the two rays as shown in Figure 6.40b.

Draw on the normals for the incident beam and the exit beam. Now measure the angle of incidence and angle of refraction. Use Snell's law to find the refractive index of the block.

Apparent depth

Figure 6.41 shows a small object 'O' under water. If it is a light bulb, it is giving off its own light; if it is the tip of a fish's tail, the light is being reflected off it. Light within the cone drawn in the diagram ends up in your eye. Be sure you understand that the refraction as shown is correct: light coming out from the water will be bent away from the normal.

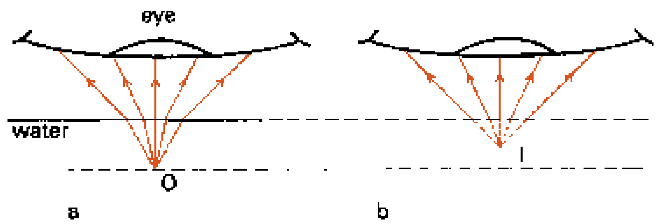


Figure 6.41 *The light, which starts from 'O', enters the eye as if it was spreading out from 'I'.*

What does the eye – or, more accurately, the brain – register? Look at Figure 6.41b. If no water was present but instead the object was placed at point 'I', the light reaching the eye would be exactly the same. To realise this, cover up the bottom part of diagrams (a) and (b) with a sheet of paper so all you can see is the eye and the light entering it, and you will find the two are identical.

Combining the two in a single diagram, and showing only the edges of the cone of light, we get Figure 6.42. What is really there is the object at 'O'. What we see is the image at 'I', and this appears to be nearer the surface. Notice that the lines from the surface to point 'I' are shown dotted. This is because no light takes that part of the path: they are just construction lines to pinpoint where the image appears to be.

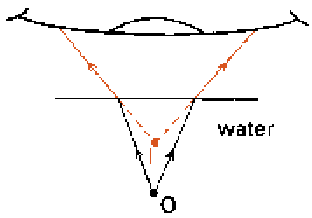


Figure 6.42 *The apparent depth of the image 'I' is less than the real depth of the object 'O'.*

It is possible to calculate just how marked this effect will be, and relate it to the refractive index n of the liquid. The relationship between apparent and real depths is given by:

$$\frac{\textbf{real depth}}{\textbf{apparent depth}} = n$$

This is not a definition of refractive index, merely a way of predicting what the apparent depth will be. Nevertheless, it offers a way to find the refractive index of the liquid.

Activity 6.14: Using refractive index = $\frac{\text{real depth}}{\text{apparent depth}}$ to find the refractive index of a liquid

Place the end of a ruler into some water in a tank, holding it at an angle to the surface. View it from above. Mark the real depth A and the apparent depth A′ on the side of the tank. Measure real and apparent depth. Use $\frac{\text{real depth}}{\text{apparent depth}} = n$ to estimate the refractive index of the liquid. Repeat with other liquids.

Why does the ruler appear to be bent?

The effect is due to the refraction of light.

Indeed, it is another example of apparent depth. Looking at the ruler from above, the end A will seem to be at A′. Similarly B and C look as though they are at B′ and C′ respectively, and so the ruler seems to be bent.

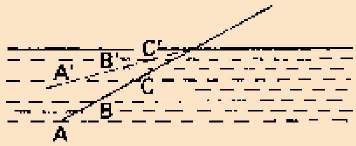


Figure 6.43

Total internal reflection

Suppose you find yourself under water with a torch that gives a narrow beam. Figure 6.44 illustrates the effect of shining the light on to the underside of the water surface at a progressively larger and larger angle (as measured to the normal, remember). Only the path of the main beam is shown – there will always be some reflection back as well.

At a particular angle of incidence (marked *c*) the light emerges along the surface.

What will happen if you shine the light at a greater angle of incidence than this?

Where will the beam come out? Activity 6.16 explores this.

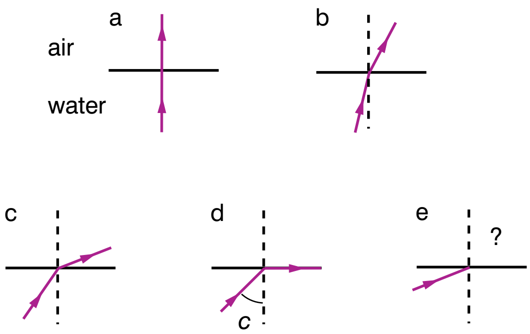


Figure 6.44 *What happens as the angle of incidence increases?*

Activity 6.15: Observations that indicate that light can be refracted

In a small group, make a list of observations that indicate that light can be refracted (you could start with the observations from the activities you have already carried out in this section).

Explaining the results

The light in the last section of Figure 6.44 does not emerge. Despite the fact that the air is a transparent medium, no light enters it.

What happens to the energy of the incident light beam? There has always been a second possibility, and that was the partial reflection back from the surface. Now there is only that one possibility. The reflection back is total: for light hitting it at such an angle, the inside of the water surface behaves as a perfect mirror.

Activity 6.16: Exploring increasing the angle of incidence

Figure 6.45 shows one possible arrangement.

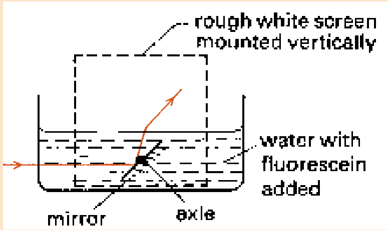


Figure 6.45 *Where does the light come out?*

The bright beam of light may be obtained from a slide projector with a horizontal slit in place of a slide. The mirror can be glued to an axle mounted in a rubber 'sucker' pressed against the side of the tank; to alter the angle at which the beam of light is incident on the underneath of the water surface, rotate the mirror.

Add fluorescein, a yellow-green dye which glows brightly along the path of the light to make it visible to the water. The route the light takes after emerging from the water can be traced on a rough white screen.

KEY WORDS

total internal reflection *occurs when light strikes a medium boundary at an angle of incidence greater than the critical angle, and all light is reflected*

critical angle *the angle of incidence on a boundary above which total internal reflection occurs*

This is called **total internal reflection**. It will occur only if two conditions are fulfilled: 1) that the light is inside the water or the glass, trying to 'escape'; 2) that the light hits the inside of the surface at a sufficiently large angle to the normal.

This may be stated more formally. Total internal reflection will occur if:

1. **Light travelling in a medium such as water or glass comes to a surface with a medium in which it travels faster (usually air).**
2. **It hits the inside of this surface at an angle of incidence greater than the critical angle.**

The **critical angle** is defined as:

the particular angle of incidence for which the light emerges along the surface (at an angle of refraction of 90°).

The critical angle is the angle marked as *c* back in Figure 6.44.

Worked example 6.6

Ordinary glass has a refractive index of about 1.5. Work out its critical angle at a boundary with air.

There are two points to bear in mind here: that at the critical angle refraction still takes place (just) as in Figure 6.46, and that sin 90° = 1.

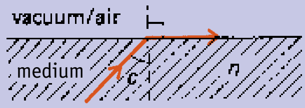


Figure 6.46 *At the critical angle, the angle of refraction is 90°.*

Using the relationship $n = \dfrac{\sin\theta_1}{\sin\theta_2}$ we get:

$n = \dfrac{\sin 90}{\sin c}$, but as sin 90 = 1, $\sin c = \dfrac{1}{n}$

The refractive index of ordinary glass is about 1.5 and its critical angle at a boundary with air can be worked out from:

$\sin c = \dfrac{1}{n} = \dfrac{1}{1.5} = 0.667$

so $c = 42°$

Total internal reflection and its use in optical fibres

If the glass block you considered in the last section becomes elongated into a rod, it will work in just the same way as it did before. Once introduced into one end, the light cannot escape so long as it always hits the side walls at an angle of incidence greater than the critical angle for the material of the rod. The light is trapped in what seems like a kind of pipe with silvered walls, as shown in Figure 6.47.

Recent advances in technology have led to a whole range of applications of this effect. Individual plastic fibres can be made which, optically insulated from one another, may be enclosed side by side in a non-transparent sheath. The result is a bundle no wider than the average electrical wiring to lamps, and so flexible that it can be tied in knots without affecting its performance.

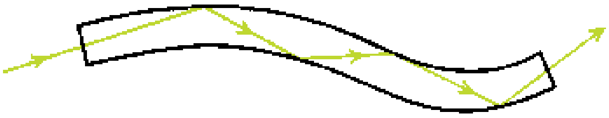


Figure 6.47 *The principle of fibre optics.*

Applications of fibre optics

An endoscope is a device used by doctors to see inside the body. It consists of two bundles of plastic fibres which can be passed down the throat, for example, to view the stomach. One bundle takes the light down to illuminate the area, while the other takes the reflected light back to construct the image.

Telecommunication companies can use light waves to convey messages and information in much the same way as they use radio waves. The possibility is very tempting because, for technical reasons, a single light beam can transmit far more telephone conversations at one time than a radio wave. Using a light beam raises severe problems, however. It does not travel well through the atmosphere, and on a foggy day communications would be cut completely. There is also a lot of synthetic interference which might swamp the signal, from street lights, bonfires and the like.

Fibre optics can provide the answer. Nowadays, in some places, a single cheap fibre optic cable has replaced several telephone wires. The material of the fibres has been made fantastically transparent, so the light can be detected at the far end of a cable as much as 20 km long.

A fibre can be used to monitor the temperature inside a jet engine. The optical cable carries the radiation from the hot surfaces inside to a pyrometer mounted outside the engine.

Convex and concave lenses

Before we proceed further, we must sort out a few points. The first of these is how to spell! Notice that in the singular 'lens' does not have an 'e' on the end. It is really a Latin word, and means 'bean'. Early lenses were just little blobs of glass and people were struck by their similarity in shape to beans, hence the name.

Figure 6.49 shows a **convex lens** acting on a beam of parallel light to bring it to a focus. Like a prism, it works by refraction at the two faces, but for convenience we think of the lens as bending the light once as shown.

The principal axis of the lens is the line passing through the centre of the lens, perpendicular to it.

Activity 6.17: Transmission of light through a fibre optic cable

Given a length of optical fibre cable, design a demonstration to show how it will transmit light.

As well as the cable you will have a white light source, a triangular glass prism and a suitable Perspex screen (either made of milky Perspex or with a roughened surface). If you feel you will need additional apparatus, consult your teacher.

Figure 6.48 *Fibre optics can also be used to produce decorative lighting such as this lamp.*

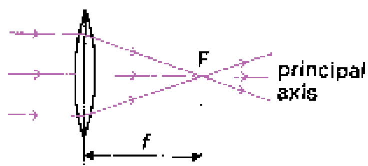


Figure 6.49 *A convex lens focusing a beam of parallel light.*

KEY WORDS

fibre optics *glass or plastic fibres that carry light along their length*

concave lens *a lens having at least one surface curved like the inner surface of a sphere*

dioptre *a unit of measurement of the optical power of a lens or curved mirror*

convex lens *a lens having at least one surface curved like the outer surface of a sphere*

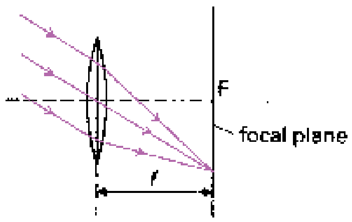


***Figure 6.50** Any parallel beam of light is brought to a focus somewhere in the focal plane.*

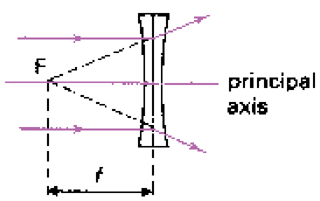


***Figure 6.51** A concave lens.*

Light which enters the lens parallel to its principal axis will be converged to the point marked 'F', the principal focus. The distance '*f*' from the centre of the lens to point F is called the focal length of the lens, and is a measure of the strength of the lens: a strong lens will have a short focal length. Since a lens is reversible there will be a point F on each side of the lens, equal distances from it.

An alternative way to measure the strength of a lens, used by opticians, is to specify its power *p*.

The power of a lens = $\frac{1}{\text{(its focal length in metres)}}$,

so the larger numbers go with the stronger lenses. The unit of power in the optical sense will be m^{-1}, given the name **dioptres**. Thus a lens whose focal length is 20 cm would have a power of $\frac{1}{0.2}$ m = 5 dioptres.

The focal plane of the lens is where the surface of a screen would lie if it was placed the focal length away from the lens as shown in Figure 6.50. A parallel beam of light which does not enter along the principal axis will be brought to a focus somewhere in the focal plane as in the diagram.

A **concave lens** is similar, except that a beam of light parallel to its principal axis will emerge spreading out, as if coming from its principal focus as shown in Figure 6.51.

Activity 6.18: Comparing lenses and mirrors

Compare the terms used for lenses with those used for mirrors (see page 205). Draw a diagram of a concave lens and a convex lens and mark on the diagrams: principal focus, principal axis, focal point and radius of curvature.

Magnification

This is just a number: the number of times the image is larger than the object. A magnification of less than 1 means that the image is smaller.

Whenever we refer to magnification we mean linear magnification, defined as the height of the image divided by the height of the object. Two examples are shown in Figure 6.52.

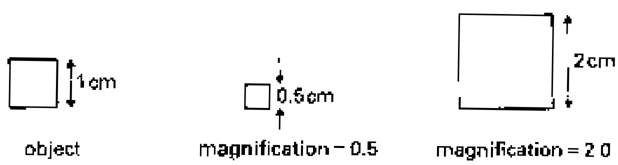


Figure 6.52

Notice that when the magnification is 2, the area of the square has increased fourfold. Some advertisements for binoculars give their area magnification: an area magnification of 49 sounds much better than a magnification of 7, but they are the same thing!

Activity 6.19: Using a convwx lens

Try getting a convex lens to cast an image on a screen. Where must the screen be placed? How big is the image? Is it the right way up or inverted?

Figure 6.53 shows one possible set-up. The object is a transparent plastic ruler with clear centimetre and millimetre markings held in a retort stand and clamp.

Start with the object as far back from the lens as you conveniently can, then adjust the screen until the image on it is in sharp focus. Measuring distances from the centre of the lens, record in a table the object distance (to the ruler) and the image distance (to the screen). Note also the magnification of the image, from the spacing of the centimetre markings on the screen and whether it is inverted or the right way up.

Keep moving the object closer to the lens in stages, and each time repeat the measurements.

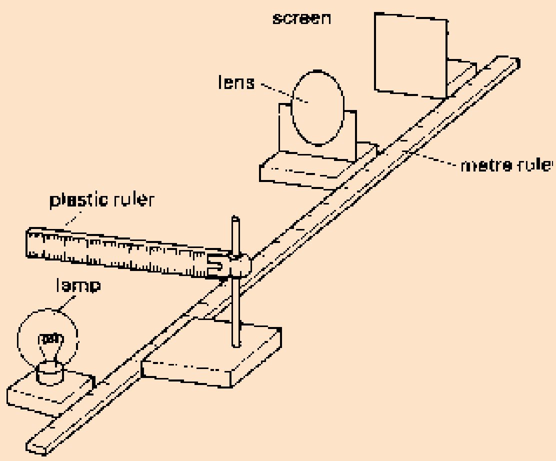


***Figure 6.53** Using a converging lens to cast an image on a screen.*

The thin lens formula

The formula that you met in the last section with curved mirrors also applies to lenses. To remind you, the notation used is this:

f = the focal length of the lens

s_o = the distance from the object to the centre of the lens

s_i = the distance from the centre of the lens to where the image is formed

The connection between them is:

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f}$$

To represent a concave lens, we give the focal length a negative value. In a calculation, if the image distances works out to be negative, that tells us that the image is not a real one but a virtual one.

You can also predict the magnification from:

$$\textbf{magnification} = \frac{s_i}{s_o}$$

Finding the position and nature of an image formed by a convex and concave lens

To be able to say how big an image will be and which way up it is, we need to take an object which has a definite size and which has a top and bottom to it. The simplest object to choose is an arrow, as shown overleaf in Figure 6.54.

Worked example 6.7

An object is placed 40 cm from a converging lens of focal length 20 cm. Where will the image be formed, and what is its magnification?

s_o = 40 cm = 0.4 m and f = 20 cm = 0.2 m.

Putting these values into

$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f}$ gives us:

$$\frac{1}{0.4} + \frac{1}{s_i} = \frac{1}{0.2}$$

so $2.5 + \frac{1}{s_i} = 5.0$.

This means that:

$$\frac{1}{s_i} = 5.0 - 2.5 = 2.5$$

so $s_i = \frac{1}{2.5} = 0.4$ m (40 cm).

The magnification =

$$\frac{s_i}{s_o} = \frac{0.4}{0.4} = 1.0$$

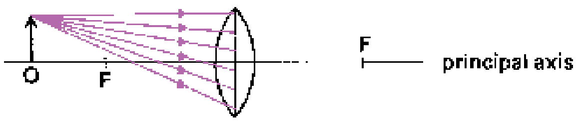


Figure 6.54 *A cone of light will reach the lens from the tip of the arrow.*

The foot of the arrow is placed on the principal axis, and we already know that light spreading out from a point on the principal axis will be brought together by the lens somewhere further along it.

We do not consider the image of the bottom of the arrow, therefore, but we concentrate instead on the light spreading out from its tip. A cone of light will reach the lens, and the focusing action of the lens will bring this cone of light together again to a point. Our problem is to say where this point will be.

Luckily within that cone there are two particular directions of travel for which we can predict the path of the light as it leaves the lens. The two special cases are these:

1. Light entering the lens parallel to its principal axis will be refracted through the focus 'F', as shown in Figure 6.55.
2. Light passing through the centre of the lens will carry on undeviated, as in Figure 6.56.

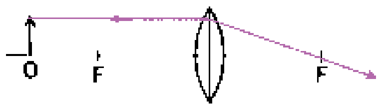


Figure 6.55 *Light parallel to the principal axis will be refracted through the focus.*

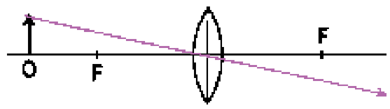


Figure 6.56 *Light passing through the centre of the lens will carry on undeviated.*

The second case needs a word of explanation. If a lens can be thought of as a series of prisms, then the middle section is a rectangular glass block. Light passing through such a block at an angle emerges still going in the same direction but shifted sideways, as shown in Figure 6.57.

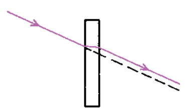


Figure 6.57 *The sideways displacement is negligible.*

A lens is very thin, perhaps 5 mm or thereabouts, and therefore the extent of the sideways displacement is so slight as to be negligible.

Figure 6.58 shows both these special 'predictable' rays combined on one drawing. Where these two lines meet at a point, so too will the rest of that cone of light from the tip of the object.

You can now draw the complete ray diagram for a converging lens being used as a projector. O is the slide, placed just outside the focus of the lens and strongly illuminated from behind; I is the enlarged real image formed on a suitably positioned distant screen (Figure 6.59).

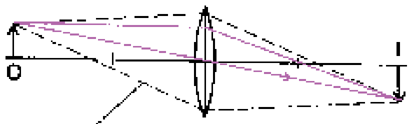


Figure 6.58 *Where the cone of rays shown in Figure 6.54 will meet.*

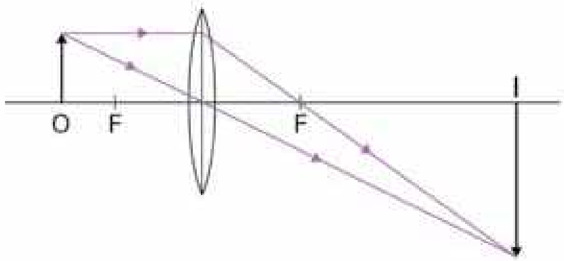


Figure 6.59 *Ray diagram for a converging lens being used as a projector.*

In the ray diagram for a camera, O is the distant object being photographed, and I is a small real image just beyond the focus of the lens, which is cast on light-sensitive film at the back of the camera (Figure 6.60).

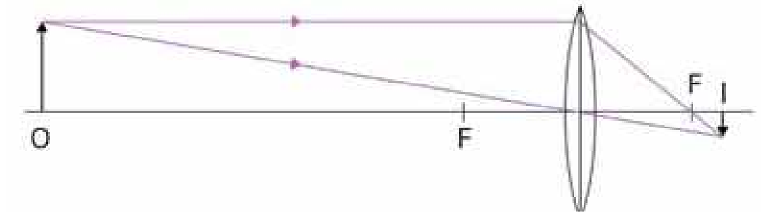


Figure 6.60 The ray diagram for a camera.

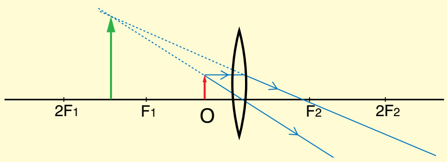
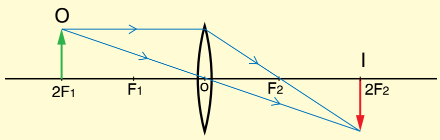
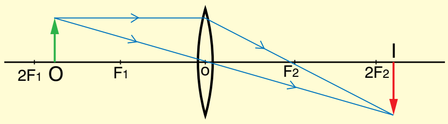
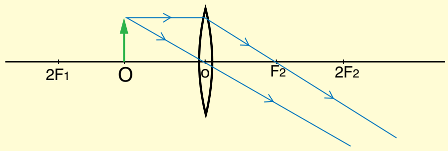
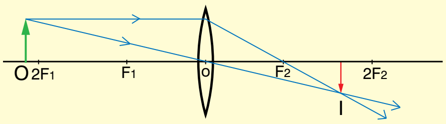
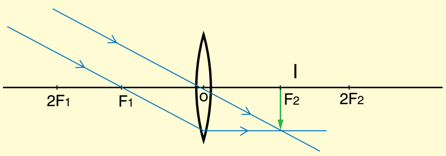
Position of object	Key diagram	Image is
Between F_1 and lens		virtual erect magnified
At $2F_1$		real inverted same size as object
Between F_1 and $2F_1$		real inverted magnified
At F_1		formed at infinity real inverted magnified
Beyond $2F_1$		formed between F_1 and $2F_1$ real inverted diminished
At infinity		formed at F_1 inverted real highly diminished

Figure 6.61 How to find the position and nature of the image formed by a convex lens using a ray tracing method.

How an image is formed due to combination of thin lenses

To view an image and make best use of the available light, you need an arrangement such as that shown in Figure 6.62.

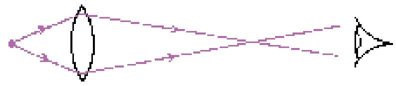


Figure 6.62 *One method of viewing an image.*

It is no good putting your eye at the point 'I' itself – that would be like trying to read this print by resting your eyeball on the page! You must stand back so you can focus on it clearly.

The first stage in forming an image with a combination of thin lenses can therefore be drawn as in Figure 6.63 (though the angles are much exaggerated for clarity).

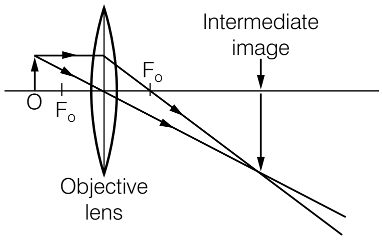


Figure 6.63 *The first stage in forming an image.*

The image, labelled 'intermediate image' in Figure 6.63, is typically 10–15 cm back from the objective.

In principle, you could cast this image onto a screen put there, but in practice it would be far too faint. It would be better to try to view it by placing your eye back to the right of the drawing – it does not matter how far back, so long as the image is outside your near point so you can focus on it.

In a combination of thin lenses you achieve a second stage of magnification by looking at the image not directly but through a second lens.

The intermediate image must therefore lie inside the focus F_e of the second lens. Unfortunately the two rays whose progress we have followed so far are not 'special' ones for the second lens – they will be refracted through it and help to form the virtual final image which we see, but their path is not predictable. Therefore we have had to add some construction lines (shown in blue) to see where the image would be produced, and then draw the rays emerging from the second lens spreading out from there (Figure 6.64).

KEY WORDS

objective lens *first lens encountered by incoming light rays from the object*

eyepiece lens *lens nearest to the eye*

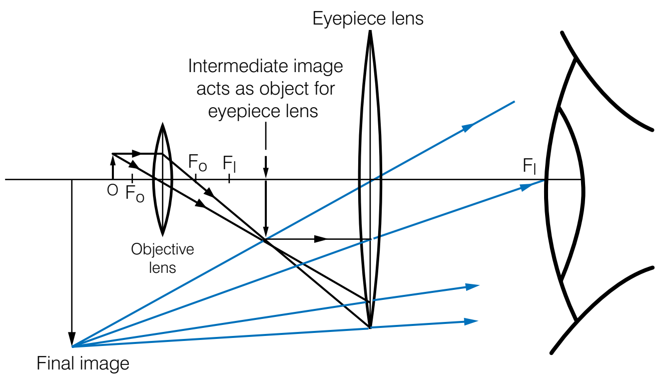


Figure 6.64

The simple microscope

A simple microscope is a magnifying glass. If the object lies between the focus of the lens and the lens itself, the light comes out diverging, as shown in Figure 6.65.

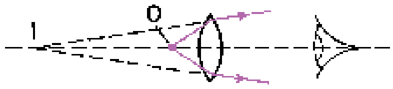


Figure 6.65 You will see a virtual image at 'I'.

No real image will be formed: wherever you put a screen, you cannot recreate on it an image of the original point of light. As you move the screen back from the lens, all you can get is an ever-widening circle of light.

If you place your eye as shown, however, you should see something: a virtual image. Light enters the eye as if it was coming from the point I, located somewhere behind the lens and further away than the object really is.

The image you see is the right way up and enlarged, as shown in Figure 6.66.

Figure 6.66 Virtual image the right way up and enlarged.

If we trace the progress of the two 'predictable' rays, this time they emerge from the lens still spreading out. No real image will be produced, but if you put your eye as shown in Figure 6.67, you will be able to see a virtual image located back behind the lens at I.

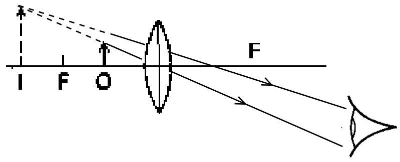


Figure 6.67 The image you see is upright, enlarged and further back than the object.

The telescope

A telescope is designed for seeing more detail in an object that is a long distance away.

If we are using a lens, the only magnified images that are possible are formed with a converging lens having the object within 2 f (as in the projector) or within f (as in the simple microscope). Neither is possible if we are looking at the surface of the moon, so how do we do it?

The solution is to use the objective lens to form a real image in its focal plane. You then examine that image through a magnifying glass (the eye lens) or a whole set of converging lenses (the eyepiece).

Let us consider each stage in turn. The crucial factor in a quality telescope is the objective. What we examine is the image it produces, not the real object, and if the fine detail is not present in that image, we will never be able to see it. It would be like looking at a photograph in a newspaper through a microscope: all we would see is a big blur because no further detail is present in it. Therefore a telescope is often described just in terms of its objective – what sort is being used, and what its diameter is.

The image produced by the objective is, of course, smaller than the original object – not desirable, but unavoidable. If you look at the images of the window frames through a series of different lenses, you will find that the stronger lenses form smaller images.

Here therefore we need a weak lens, so the image is still as large as possible – the main limit is the length of the instrument.

The eye lens that we use to look at that image needs to be a strong one, as with the microscope. If we set it so the final image is offered to us back at infinity, that means having the intermediate image at its focus. Therefore the total length of the telescope will be the focal length of the objective, f_o (which will be large), added to the focal length of the eye lens, f_e (which is tiny) (Figure 6.68).

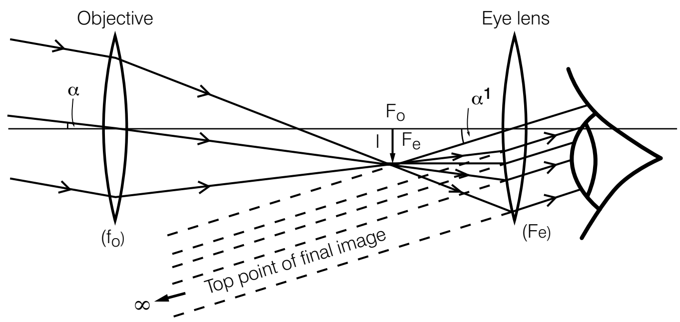


Figure 6.68 *Telescope.*

The light rays traced through the telescope started from one off-axis point back at infinity. The two angles marked with α indicate the angles subtended with and without the telescope. Their ratio is the angular magnification.

It is possible to work out the angular magnification of the telescope used like this by the ratio $\frac{f_o}{f_e}$.

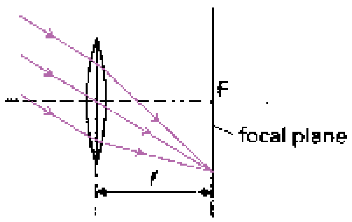


Figure 6.69

The camera

If the lens and screen in Figure 6.69 are enclosed in a light-tight box, you have a camera.

For distant objects a small upside-down picture of the outside world will be cast on the film, which really just consists of light-sensitive chemicals mounted in the focal plane of the lens. In a digital camera, the film is replaced by an image capture surface that converts the image into a computer file.

If the object being photographed comes in much closer, then the distance from the lens to the film must increase slightly. With most modern cameras this is done by rotating the lens: a screw thread then winds it a small distance backwards or forwards. The other possibility is to slide the lens along, in which case the camera is kept light-tight by a concertina-like cloth 'bellows'.

The human eye

The human eye (see Figure 6.70), works rather like that of the camera, where a converging lens forms a tiny upside-down image of the distant world on a screen. The lens in the eye is not a glass one, of course, but is made of living tissue. The screen on which the image is cast is not a white board; instead, it is the back surface of the eyeball (called the **retina**), packed with light-sensitive cells all connected straight to the brain via many separate nerves, which leave the eye in a bundle called the **optic nerve**.

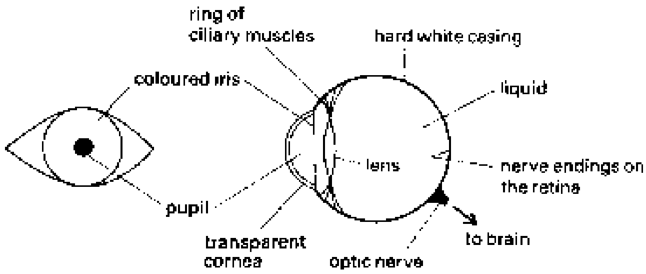


Figure 6.70 *The structure of the eye.*

Each eye is enclosed in a tough white casing whose front part is visible as the white of your eye. The bulge at the front of the eyeball is the transparent window (the cornea) by which the light enters. The amount of light reaching the retina is controlled by a ring called the iris – usually brown or blue. It is the iris that determines the colour of your eyes.

The size of the hole in the middle of the iris (the **pupil**) adjusts itself as the brightness of the light changes. In bright light the pupil is just a dot, while at night the iris opens right up to leave a large pupil. The pupil looks black because the inside of the eye is covered in a black pigment. The effect of this is to absorb stray light, and so prevent it from being reflected onto the retina (otherwise the contrast in the image would be reduced, rather like using a slide projector in a poorly darkened room).

The lens in the eye is surrounded by a circular sheet of muscle, which divides the eyeball into two quite separate regions. The eye is filled with liquid, which gives rigidity to the whole structure. The liquid in the front compartment is salty, rather like tears, while the space between the lens and the retina contains a thicker, less runny liquid. Tear drops themselves are produced by glands in the sockets into which the eyeballs fit, and act as a lubricant for the outside surface of your eyes.

How the eye focuses

You can look at close-up objects or distant ones. Unless you require spectacles, the image in both cases is clearly focused on the retina.

A camera would achieve this by altering the position of the screen (the film) relative to the lens, by moving the lens slightly. With the eye the story is quite different, however. Most of the refraction of the light, to make it converge to a focus on the eye's retina, takes place as the light first enters the eyeball at the boundary between air and the cornea. This is where the greatest change in the speed of light occurs, and it is that speed change which causes the refraction.

The fine control of the focusing is done by the lens, but it must be realised that only a little of the total refraction takes place there. As the light goes between the liquid in the eye and the lens, there is only a small change in speed and therefore only a small difference in refractive index.

KEY WORDS

retina *light-sensitive cells lining the inner surface of the eye*

optic nerve *a nerve transmitting visual information from the retina to the brain*

pupil *an opening in the centre of the iris of the eye*

KEY WORDS

accommodation *the eye's ability to focus on objects at various distances*

Activity 6.20: Exploring 'near points'

Your 'near point' is the closest distance at which you can focus clearly. Bring this book up towards your eye and eventually the print becomes blurred. The lens in your eye has then reached the limit of its adjustment; it will bulge up no more.

Locate your near point in this way, and get a friend to measure how far it is from your eye.

Carry out a survey of other people's near point distance. Include people much older than you. If someone wears spectacles, take the measurement both with and without them: do the spectacles make any difference?

It is said that as people age the lens in their eye becomes less supple and so their near point gets further away. Can you find any evidence to support this in your survey?

Activity 6.21: Compare and contrast the structure and functions of the human eye and the camera

In a small group, use the information you have learnt so far to draw up a comparison between the camera and the human eye.

The eye's ability to focus on objects at varying distances is called **accommodation**. It is the ring of muscles around the lens that enables the eye to accommodate. With a distant point object the light is almost parallel, and the eye's lens focuses it back to a point on the retina. When you view a near object, the light spreading out from it and reaching your eye will be diverging strongly and yet the same lens has to focus it on the same retina. A stronger lens is needed to accomplish this second feat, and the muscles achieve this by causing the lens to bulge up into a fatter more rounded shape.

Figure 6.71 illustrates this process in a well-adjusted eye which does not need spectacles. For simplicity, all the bending of the light is shown occurring at the lens.

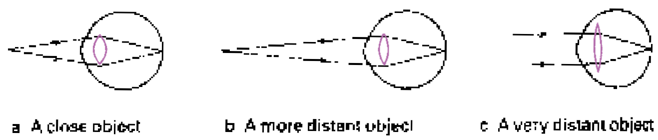


Figure 6.71 *The lens changing shape to focus light from objects at differing distances.*

In Figure 6.71a the eye is focusing on a close object. The lens has to be made into a very rounded shape to refract the light enough for it to meet on the retina. In Figure 6.71b the object is more distant. The light is not diverging so much, so less bending is needed to make the rays meet on the retina, and the lens is not so 'bulged up'. Finally, in Figure 6.71c the object is very distant. The apparently parallel rays are actually diverging from a point a long way off. Less refraction is needed and the lens is its natural shape and the muscles are relaxed.

Defects of the eye and their correction with lenses

The lens in the eye needs to be at its weakest when you are viewing objects a long way away, at infinity. With a correctly adjusted eye, at this point your lens is completely relaxed.

As an object approaches, the power of your eyeball has to increase, to refract the light a greater amount so it still focuses on the retina. To accomplish this the lens bulges into a fatter shape, thus giving it a shorter focal length. There is a limit to how much your lens can bulge, and when this limit is reached the object is at your near point. With a correctly adjusted eye the near point is taken to be 250 mm, though many young people can manage shorter distances than that.

The first three defects in vision described below all involve the lens in the eye. The fourth one is due to the shape of the cornea.

1 *Short sight* (*myopia*). This happens if the lens is too strong for the eye or, looked at another way, the eyeball is too long for the lens. With an object at the far point (that is, the greatest distance which can be focused clearly), the lens is fully relaxed – and for this eye the far point is not all that far away!

For objects at greater distances the lens can go no weaker, so light from them is made to meet in front of the retina and so the image is blurred.

The one compensation is that the near point will be exceptionally close.

To correct this fault, a diverging lens must be placed in front of the eye so parallel light is made to enter the eye as if it was spreading out from the eye's far point (see Figure 6.72).

2 *Long sight* (*hypermetropia*). This time the lens is too weak. The parallel light from distant objects would not be focused by the relaxed lens until past the retina, but they can still be seen clearly by causing the lens to bulge – thus using up some of the available accommodation already. This means that as an object approaches, the lens soon bulges to its maximum extent. Thus the near point will be an inconveniently large distance away.

This fault may be corrected with spectacles containing converging lenses, to strengthen the eyeball's optical system (see Figure 6.72).

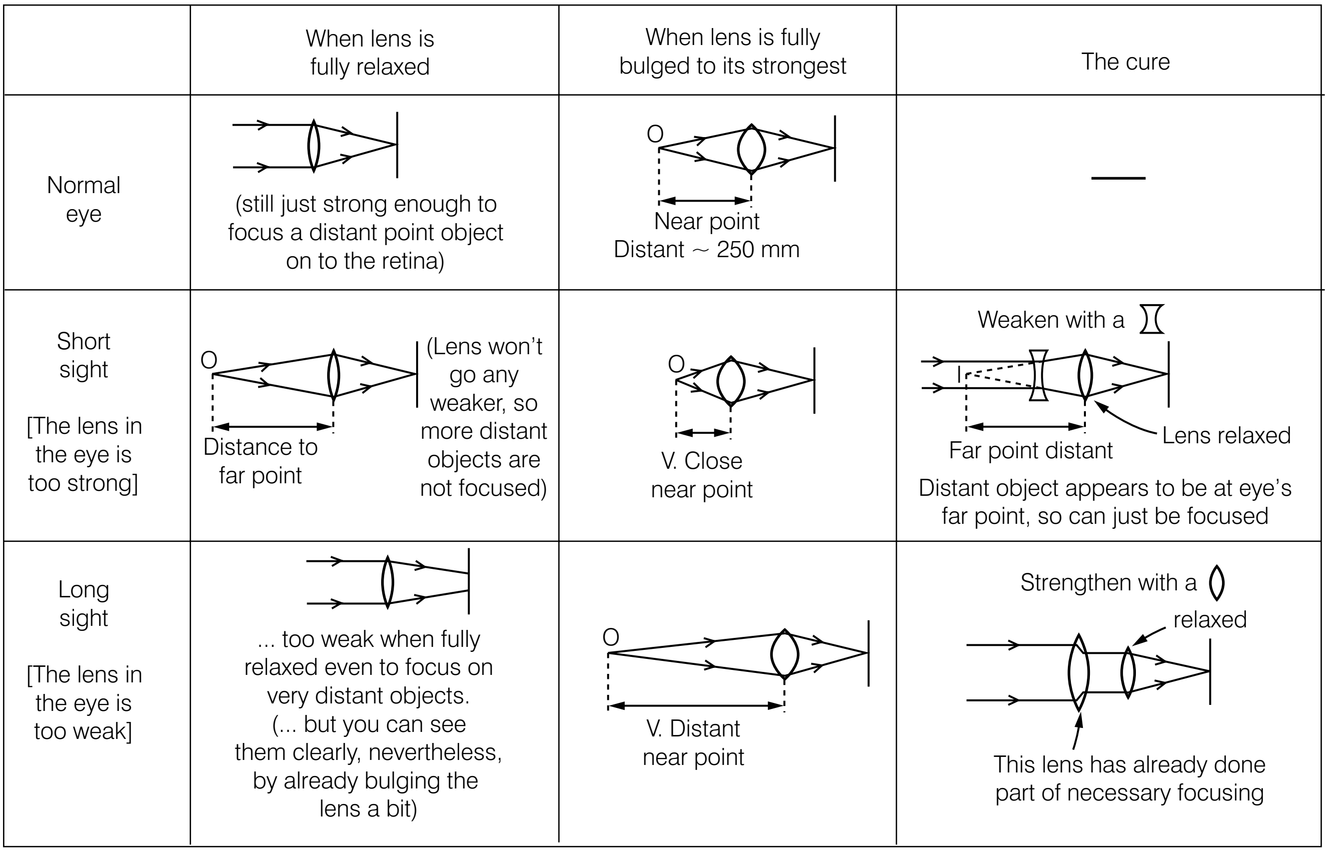


***Figure 6.72** Sight defects.*

3 *Old sight* (*presbyopia*). As people age, the lens in their eye may become less supple. In that case the power of accommodation may become affected at both ends of the range – their near point is too far away, so a book has to be held at arm's length, and their far point is too close so they cannot see distant things clearly.

In that case a pair of reading spectacles with diverging lenses and a pair of general viewing spectacles with converging lenses may be needed, or else a single pair of bifocals.

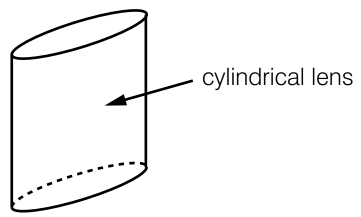


Figure 6.73 *Cylindrical lenses.*

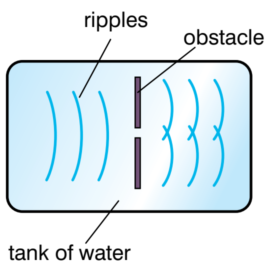


Figure 6.74 *Diffraction.*

KEY WORDS

diffraction *the change of direction of a wave at the edge of an obstacle in its path*

4 *Astigmatism*. This problem arises if a person's cornea has a different curvature in the horizontal plane from that in the vertical plane. This results in two slightly different powers. Vertical lines in the field of view may be sharply focused, for instance, while horizontal lines are a bit blurred.

The remedy is a pair of spectacles fitted with cylindrical lenses, whose surfaces are each part of a cylinder rather than a sphere (see Figure 6.73). These increase the power of the eye in one plane, to bring it up to the power in the other plane.

Diffraction of light

Diffraction of light (or any type of wave) is a change in direction that happens as the waves move through or round obstacles. You can see this effect with water waves if you place an obstacle in a tank of water and cause waves of different wavelengths to go through the obstacle (see Figure 6.74).

Diffraction is greatest when the wavelength of the waves is long compared to the size of the gap or obstacle.

Activity 6.22: Diffraction of light

Use two pencils or other straight edges. Place tape around the shaft of one pencil to create a space between them, as shown in Figure 6.75.

Darken the room and then look through the slit between the pencils and observe a candle flame at a distance of about 2 m (take care with the candle!). What do you see? Rotate the pencils and describe how your observations change. Change the width of the gap between the pencils slightly. What do you see now?

Figure 6.75 *Diffraction of light.*

Activity 6.23: Two-slit diffraction

Set up the apparatus as shown in Figure 6.76. What do you see?

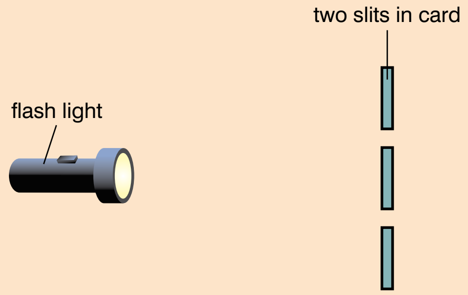


Figure 6.76 *Apparatus for two-slit diffraction.*

Activity 6.24: Using a diffraction grating

Set up the apparatus as shown in Figure 6.77. What do you see?

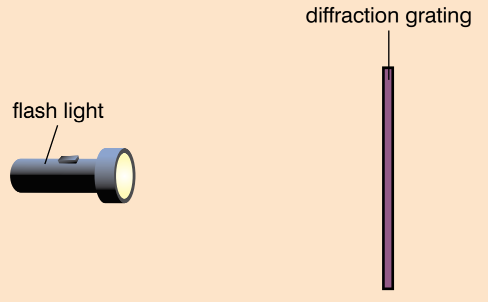


Figure 6.77 Apparatus for a diffraction grating.

In a small group, look back to the section on the electromagnetic spectrum on page 198 and see if it helps you to explain your observations before you read on.

KEY WORDS

diffraction grating *a material with a large number of narrow, regularly spaced slits, designed to produce a diffraction pattern*

Explaining the dispersion of white light to produce a spectrum

White light has a range of wavelengths, from blue to red. Since wavelength and speed are related by the equation $v = f\lambda$ (see page 197), then a range of wavelengths will produce a range of speeds. You know from page 215 that the amount of refraction is related to speed and so the different wavelengths in white light are refracted by different amounts to produce a spectrum, as shown in Figure 6.79. The rays are deviated by the prism.

Figure 6.79 Producing a spectrum by refracting white light through prism.

Activity 6.25: Exploring dispersion of white light

Set up the apparatus as shown in Figure 6.78. What do you see?

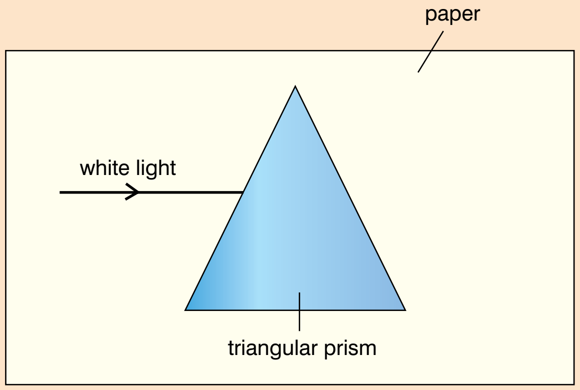


Figure 6.78 Apparatus for exploring dispersion of white light.

Activity 6.26: The CD spectroscope

Use a cracked CD to build a spectroscope. Observe the spectra of a flashlight, a regular bulb, an infrared bulb, a flashlight with a coloured filter, sunlight and a fluorescent bulb. Experiment with other light sources. Observe sunsets. Record your observations in a suitable manner.

Discuss in a small group why this works.

Warning: never look directly at the Sun.

The Fresnel lens

Fresnel lenses are thinner and lighter than conventional spherical lenses. They are made with separate sections known as Fresnel zones mounted in a frame. Fresnel lenses are used to concentrate solar light for use in solar cookers, solar forges, and solar collectors to heat water for domestic use.

KEY WORDS

solar constant *the average solar power striking the Earth's atmosphere in regions directly facing the Sun is about 1370 W/m². This is the solar constant.*

Activity 6.27: Measuring the solar constant using a Fresnel lens

Use a Fresnel lens like the one shown in Figure 6.80 to measure the heat input from the Sun.

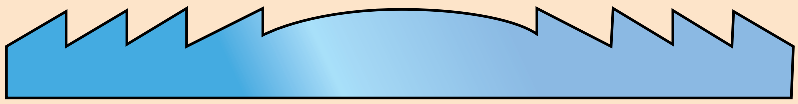


Figure 6.80 *Fresnel lens.*

Activity 6.28: Design a collector for the heat of the Sun

Design a collector that will concentrate and capture the heat of the Sun. You need to be able to heat up 5 cc of water to the highest possible temperature in 10 minutes. You may use a Fresnel lens or some other reflective surface such as mirrors or aluminium. The area of the collector must be less than 1 m².

Summary

- Refraction is the change in the direction of travel of a light beam that occurs as the light crosses the boundary between one transparent medium and another.
- The refractive index of a material is:

$$n = \frac{\sin\theta_1}{\sin\theta_2}.$$

- Snell's law can be used to solve simple problems.
- The formula

$$\text{refractive index} = \frac{\text{real depth}}{\text{apparent depth}}$$

can be used to find the refractive index of a liquid and a solid in the form of a rectangular glass block.

- A diagram representing the passage of light rays through a rectangular glass block is as shown in Figure 6.81.

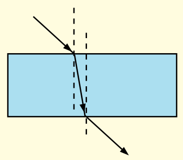


Figure 6.81

- Examples of observations that indicate that light can be refracted are the dispersion of white light to produce a spectrum and the passage of light through lenses.
- The passage of a ray of light through a parallel-sided transparent medium results in the lateral displacement of a ray.
- Total internal reflection is as shown in Figure 6.82.

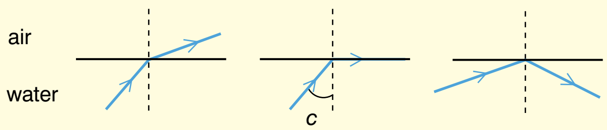


Figure 6.82

- The critical angle θ_c is the angle of incidence in a dense medium when the angle of refraction in a less dense medium is 90°.
- Total internal reflection occurs when the angle of incidence is more than the critical angle.
- You can use the formula $n = \sin 90/\sin c$ in calculations involving critical angle and total internal reflection.
- Total internal reflection is used in optical fibres because the light is trapped within the cable as shown in Figure 6.83.

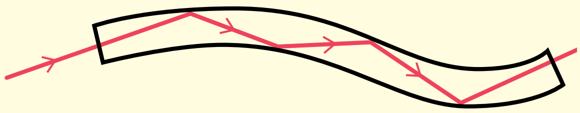


Figure 6.83

- Figure 6.84 shows the difference between convex and concave lenses and the terms used when talking about lenses.

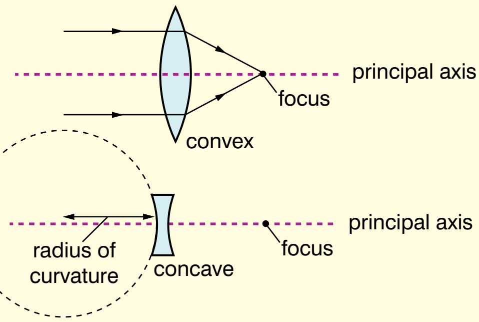


Figure 6.84

- Magnification in relation to converging and diverging lenses is $\dfrac{\text{image distance}}{\text{object distance}}$.
- When using the thin lens the equation, convex lenses have negative signs for image, object and focal length.
- The position and nature of the image formed by a convex and concave lens can be found using the thin lens formula and a ray tracing method.
- The power of a lens is defined by

$$\text{power of a lens} = \frac{1}{\text{its focal length in metres}}$$

Review questions

1. Define the term refraction.
2. Define the refractive index of a material.
3. Find the refractive index of a material where the angle of incidence is 56° and the angle of refraction is 34°.
4. Use the formula refractive index = $\frac{\text{real depth}}{\text{apparent depth}}$ to find the refractive index of a liquid in which the real depth was 5 cm and the apparent depth was 3 cm.
5. State the conditions in which refraction occurs.
6. Describe an experiment to test the laws of refraction.
7. Give examples of observations that indicate that light can be refracted.
8. Why does the passage of a ray of light through a parallel-sided transparent medium result in the lateral displacement of a ray?
9. Define the critical angle θ_c.
10. Identify the conditions necessary for total internal reflection to occur.
11. What is the critical angle for total internal reflection to occur when the refractive index of the material is 1.52 and the angle of incidence is 32°?
12. Describe how total internal reflection is used in optical fibres.
13. Distinguish between convex and concave lenses.
14. Draw a diagram to identify the meaning of: principal focus, principal axis, focal point, radius of curvature and magnification in relation to converging and diverging lenses.
15. Find the position and nature of the image formed by a convex lens using a ray tracing method.
16. Define the power of a lens.
17. Explain how the image is formed due to combination of thin lenses.
18. Draw a ray diagram to show how images are formed by lenses in a simple microscope and a simple telescope.
19. Compare and contrast the structure and functions of the human eye and the camera.
20. Describe how the human eye forms an image on the retina for different object distances.
21. Identify some defects of the eye and their correction with lenses.
22. Explain what is meant by the dispersion of white light to produce a spectrum.
23. Why does the passage of a ray of light through a triangular transparent prism result in a deviation of a ray?

End of unit questions

1. What speed do all electromagnetic waves have in a vacuum?
2. Explain some uses of electromagnetic radiation.
3. Describe an experiment to test the laws of reflection using a plane mirror.
4. The height of an object is 5 cm and the height of its image in a concave mirror is 25 cm high. What is the magnification of the image?
5. Give examples of the uses of curved (concave and convex) mirrors.
6. State Snell's law.
7. Draw a diagram representing the passage of light rays through a rectangular glass block.
8. Explain, with the aid of a diagram, what is meant by critical angle and total internal reflection.
9. Use the thin lens equation to find the focal length of a convex lens where the object is 4 cm from the lens and the image is 6 cm from the lens.
10. In December 1901 Marconi succeeded in sending the first radio signals through the atmosphere 3200 km across the Atlantic Ocean. Many scientists at the time predicted that this experiment was impossible. Their prediction would have been true for television signals. Why did Marconi's experiment work? Why was it not possible to send a television signal the same distance until the 1960s? (Hint: compare how radio and television signals are transmitted.)
11. a) State what is meant by the diffraction of waves.
    b) Draw a diagram to show how water waves are diffracted when they pass through a gap.
12. Diffraction is a property of all waves, but is only a significant effect when the wavelength of the diffracted waves is about the same size as the aperture. Explain why the diffraction of sound is easily observed in everyday life but the diffraction of light is not.
13. Calculate the critical angle for
    a) water (n = 1.33)
    b) diamond (n = 2.42)
14. A certain transparent material, A, has a higher refractive index than another material, B.
    a) Through which one does light travel more slowly?
    b) How do both speeds compare with the speed of light in air?

Index