Heinemann PHYSICS 12

4TH EDITION

VCE Units 3 & 4

Written for the VCE Physics Study Design 2017–2021

Doug Bail
Greg Moran
Keith Burrows
Rob Chapman
Ann Conibear
Carmel Fry
Mark Hamilton
David Madden
Jeffrey Stanger
Daniela Nardelli
Craig Tilley
Alistair Harkness
Jack Jurica

ALWAYS LEARNING PEARSON

Heinemann PHYSICS 12

4TH EDITION

VCE Units 3 & 4

COORDINATING AUTHORS:
Doug Bail
Greg Moran

AUTHORS:
Keith Burrows
Rob Chapman
Ann Conibear
Carmel Fry
Mark Hamilton
David Madden
Jeffrey Stanger

CONTRIBUTING AUTHORS:
Daniela Nardelli
Craig Tilley
Alistair Harkness
Jack Jurica

Pearson Australia
(a division of Pearson Australia Group Pty Ltd)
707 Collins Street, Melbourne, Victoria 3008
PO Box 23360, Melbourne, Victoria 8012
www.pearson.com.au

First published 2017 by Pearson Australia
2021 2020 2019 2018 2017
10 9 8 7 6 5 4 3 2 1

Heinemann Physics 12 4e
Series Consultant: Malcolm Parsons
Project Lead: Caroline Williams
Lead Development Editor: Vicky Chadfield
Project Managers: Susan Keogh, Hannah Turner and Shelly Wang
Editor: Sam Trafford
Series Designer: Anne Donald
Typesetter: Cam McPhail
Indexer: Max McMaster
Rights and Permissions Editors: Amirah Fatin and Samantha Russell-Tulip
Senior Publishing Services Analyst: Rob Curulli
Illustrators: Bruce Rankin and DiacriTech

Printed in China

National Library of Australia Cataloguing-in-Publication entry
Title: Heinemann physics 12 : VCE units 3 & 4 / Doug Bail [and twelve others]
Edition: 4th edition.
ISBN: 9781488611278 (paperback)
Notes: Includes index.
Previous edition: 2009.
Target Audience: For year 12 students.
Subjects: Physics—Textbooks.
Physics—Study and teaching (Secondary)—Victoria.
Dewey Number: 530

ISBN 978 1 4886 1127 8

Pearson Australia Group Pty Ltd ABN 40 004 245 943

Some of the images used in *Heinemann Physics 12* 4th Edition might have associations with deceased Indigenous Australians.
Please be aware that these images might cause sadness or distress in Aboriginal or Torres Strait Islander communities.

Heinemann
PHYSICS 12
4TH EDITION

Writing and Development Team

We are grateful to the following people for their time and expertise in contributing to the Heinemann Physics 12 project.

Malcolm Parsons
Publishing Consultant
Series Consultant

Doug Bail
Education Consultant
Coordinating Author

Greg Moran
Head of Department
Coordinating Author

Daniela Nardelli
Learning Development Leader
Subject Lead and Author

Craig Tilley
Science Writer and Assessor
Subject Lead and Author

Rachael Barker
Teacher
Reviewer

Stephen Brown
Teacher
Reviewer

Keith Burrows
Educator
Author

Rob Chapman
Educator
Author

Ann Conibear
Teacher
Author

Carmel Fry
Head of Learning Area: Science
Author

Mark Hamilton
Teacher
Author

Alistair Harkness
Teacher
Author

Jack Jurica
Teacher
Author

Melanie Lane
Head of Science
Author

David Madden
Acting Manager, QCAA
Author

Svetlana Marchouba
Laboratory Technician
Author

Dr John Nicholson
Teacher
Reviewer

Jeffrey Stanger
Head of Department
Author

Contents

Unit 3: How do fields explain motion and electricity?

Unit 4: How can two contradictory models explain both light and matter?

AREA OF STUDY 1
How can waves explain the behaviour of light?

and AREA OF STUDY 2
How are light and matter similar?

AREA OF STUDY 3: PRACTICAL INVESTIGATION

Heinemann Physics 12 4th Edition ProductLink provides extensive support material for Unit 4 Area of Study 3 Practical Investigation. This includes teacher notes and advice, logbook template and sample logbook, poster template and sample poster, rubrics, checklists and more.

How to use this book

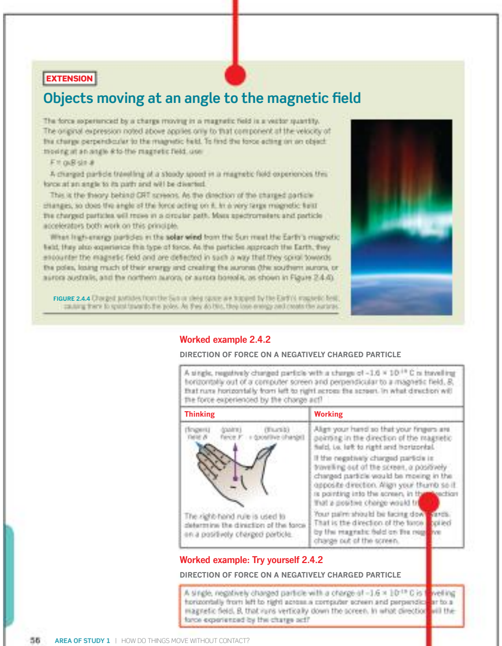
Heinemann Physics 12 4th Edition

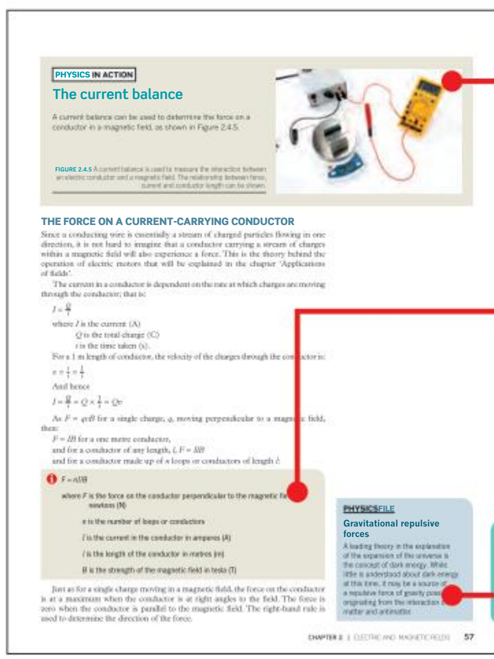
Heinemann Physics 12 4th Edition has been written to the new VCE Physics Study Design 2017–2021. The book covers Units 3 and 4 in an easy-to-use resource. Explore how to use this book below.

Chapter opener
Chapter opening pages links the Study Design to the chapter content. Key knowledge addressed in the chapter is clearly listed.

Physics in Action
Physics in Action place physics in an applied situation or relevant context. These refer to the nature and practice of physics, applications of physics and the associated issues and the historical development of concepts and ideas.

Extension
The extension boxes include material that goes beyond the core content of the Study Design and are intended for students who wish to expand their depth of understanding.





Highlight
The highlight boxes provide important information such as key definitions, formulae and summary points.

PhysicsFile
PhysicsFiles include a range of interesting information and real world examples.

Worked examples
Worked examples are set out in steps that show both thinking and working. This enhances student understanding by linking underlying logic to the relevant calculations.

Each **Worked example** is followed by a **Worked example: Try Yourself**. This mirror problem allows students to immediately test their understanding.

The fully worked solution to each **Worked example: Try Yourself** is available on *Heinemann Physics 12 4th edition ProductLink.*

Chapter review

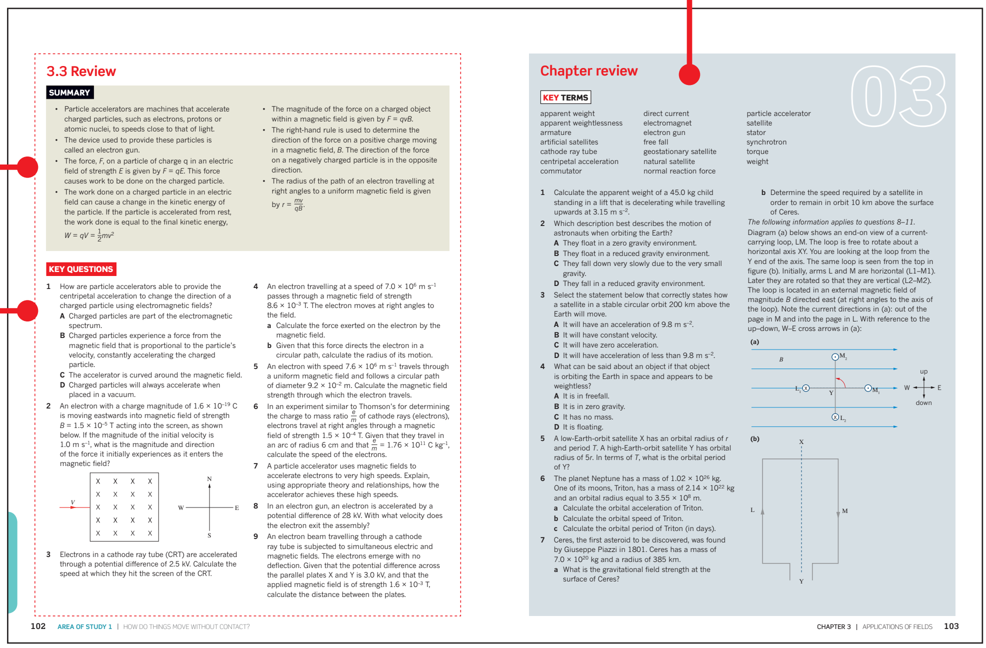
A set of higher order questions are provided at the end of each chapter to test students' ability to apply the knowledge gained from the chapter.

Section summary

A summary is provided at the end of each section to assist students consolidate key points and concepts.

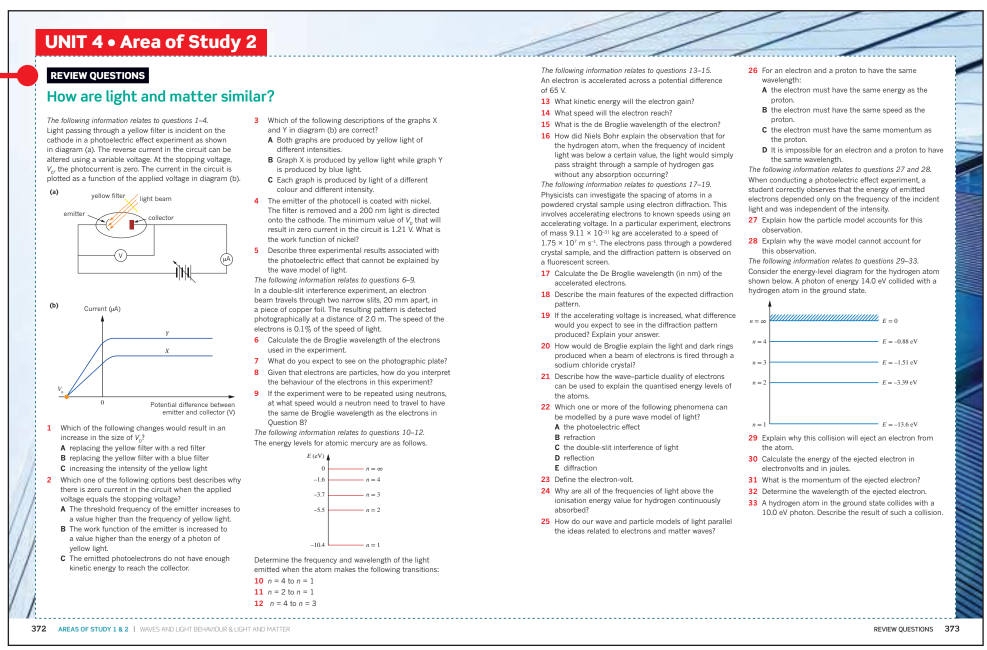
Section review

A set of 'key questions' are provided at the end of each section to test students' understanding and ability to recall the key concepts of the section as well as highlight areas that they need to revise.



Area of Study review

A comprehensive set of exam-style questions, including multiple choice and extended response are provided at the end of each Area of Study section. The questions are designed to assist students in to apply their knowledge and understanding across the entire Area of Study.



Answers

Numerical answers and key short response answers are included at the back of the book. Comprehensive answers and fully worked solutions for all section review questions, each Worked example: Try Yourself, chapter review questions and Area of Study review questions are provided via *Heinemann Physics 12 4th edition ProductLink*.

Glossary

Key terms are shown in bold and listed at the end of each chapter. A comprehensive glossary at the end of the book provides comprehensive definitions for all key terms.

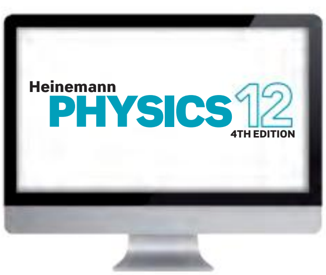
Heinemann Physics 12
4th Edition

Student Book

Heinemann Physics 12 4th Edition has been written to fully align with the *2017–2021 VCE Physics Study Design*.

Pearson eBook 3.0

Pearson eBook 3.0 is the digital version of the student book. It retains the integrity of the printed page and is available online or offline on any device—PC/Mac, Android tablet, iPad.



ProductLink

Heinemann Physics 12 4th Edition ProductLink provides comprehensive support for teachers and students completely free of charge. ProductLink includes comprehensive answers and worked solutions to all student book questions, teaching programs, risk assessments and access to SPARKlabs.

Heinemann Physics 12 4th Edition ProductLink provides extensive support material for Unit 4 Area of Study 3 Practical Investigation. This includes teacher notes and advice, logbook template and sample logbook, poster template and sample poster, rubrics, checklists and more.

UNIT 3 How do fields explain motion and electricity?

AREA OF STUDY 1

How do things move without contact?

Outcome 1: On completion of this unit the student should be able to analyse gravitational, electric and magnetic fields, and use these to explain the operation of motors and particle accelerators and the orbits of satellites.

AREA OF STUDY 2

How are fields used to move electrical energy?

Outcome 2: On completion of this unit the student should be able to analyse and evaluate an electricity generation and distribution system.

AREA OF STUDY 3

How fast can things go?

Outcome 3: On completion of this unit the student should be able to investigate motion and related energy transformations experimentally, analyse motion using Newton's laws of motion in one and two dimensions, and explain the motion of objects moving at very large speeds using Einstein's theory of special relativity.

CHAPTER

01 Gravity

Gravity is, quite literally, the force that drives the universe. It was gravity that first caused particles to coalesce into atoms, and atoms to congregate into nebulas, planets and stars. An understanding of gravity is fundamental to understanding the universe.

This chapter centres on Newton's law of universal gravitation. This will be used to predict the size of the force experienced by an object at various locations on the Earth and other planets. It will also be used to develop the idea of a gravitational field. Since the field concept is also used to describe other basic forces such as electromagnetism and the strong and weak nuclear forces, this will provide an important foundation for further study in Physics.

Key knowledge

By the end of this chapter you will have studied the physics of gravity, and will be able to:

- describe gravitation using a field model
- investigate gravitational fields including directions and shapes of fields
- investigate gravitational fields about a point mass with reference to:
  - the direction of the field
  - the use of the inverse square law to determine the magnitude of the field
  - potential energy changes (qualitative) associated with a point mass moving in the field
- analyse the use of gravitational fields to accelerate mass, including
  - gravitational field and gravitational force concepts: $g = G\frac{M}{r^2}$ and $F_g = G\frac{m_1 m_2}{r^2}$
  - potential energy changes in a uniform gravitational field: $E_g = mg\Delta h$
  - the change in gravitational potential energy from area under a force–distance graph and area under a field–distance graph multiplied by mass.

1.1 Newton's law of universal gravitation

FIGURE 1.1.1 Sir Isaac Newton was one of the most influential physicists who ever lived.

In 1687, Sir Isaac Newton (see Figure 1.1.1) published a book that changed the world. Entitled *Philosophiæ Naturalis Prinicipia Mathematica* (Mathematical Principles of Natural Philosophy), Newton's book (shown in Figure 1.1.2) used a new form of mathematics now known as calculus and outlined his famous laws of motion.

The *Principia* also introduced Newton's law of universal gravitation. This was particularly significant because, for the first time in history, it scientifically explained the motion of the planets. This led to a change in humanity's understanding of its place in the universe.

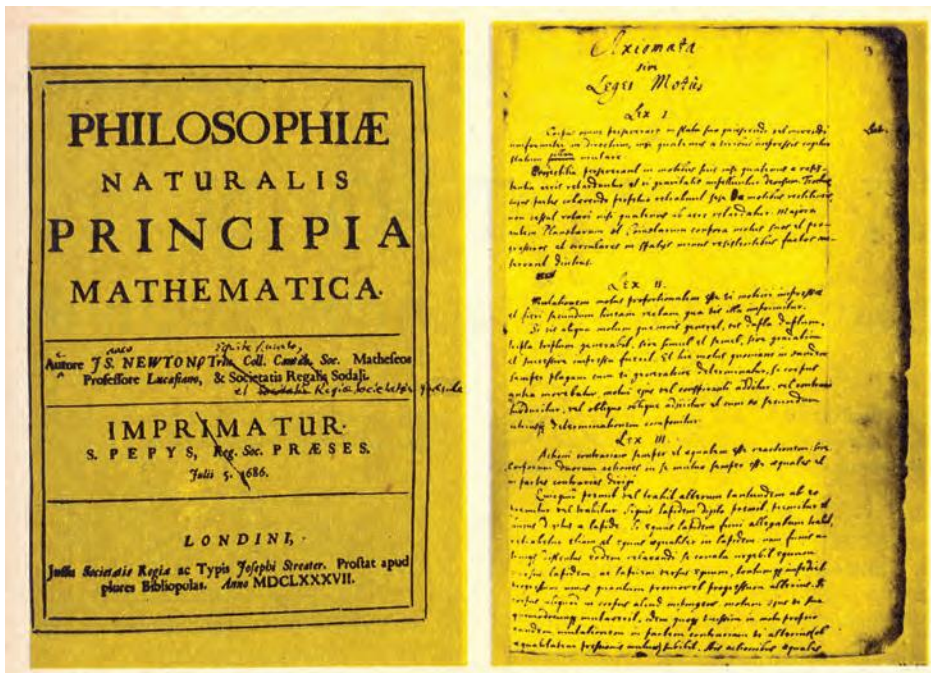


FIGURE 1.1.2 The *Principia* is one of the most influential books in the history of science.

UNIVERSAL GRAVITATION

Newton's law of universal gravitation states that any two bodies in the universe attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.

Mathematically, Newton's law of universal gravitation can be expressed as:

$$F_g = G\frac{m_1 m_2}{r^2}$$

where F_g is the **gravitational force** (N)

m_1 is the mass of object 1 (kg)

m_2 is the mass of object 2 (kg)

r is the distance between the centres of m_1 and m_2 (m)

G is the gravitational constant, 6.67×10^{-11} N m^2 kg^{-2}

The fact that r appears in the denominator of Newton's law of universal gravitation indicates an inverse relationship. Since r is also squared, this relationship is known as an **inverse square law**. The implication is that as r increases, F_g will decrease dramatically. This law will reappear again later in the chapter when gravitational fields are examined in detail.

PHYSICS IN ACTION

Measuring the gravitational constant, G

The **gravitational constant**, G, was first accurately measured by the British scientist Henry Cavendish in 1798, over a century after Newton's death. Cavendish used a **torsion balance** (shown in Figure 1.1.3), a device that can measure very small twisting forces. Cavendish's experiment could measure forces smaller than 1 μN (i.e. 10^{-6} N). He used this balance to measure the force of attraction between lead balls held a small distance apart. Once the size of the force was known for a given combination of masses at a known separation distance, a value for G could be determined.

FIGURE 1.1.3 Henry Cavendish used a torsion balance to measure the small twisting force created by the gravitational attraction of lead balls.

As its name suggests, the law of universal gravitation predicts that any two objects that have mass *will attract each other*. However, because the value of G is so small, the gravitational force between two everyday objects is too small to be noticed.

Worked example 1.1.1

GRAVITATIONAL ATTRACTION BETWEEN SMALL OBJECTS

A man with a mass of 90 kg and a woman with a mass of 75 kg have a distance of 80 cm between their centres. Calculate the force of gravitational attraction between them.

Thinking	Working
Recall the formula for Newton's law of universal gravitation.	$F_g = G\frac{m_1 m_2}{r^2}$
Identify the information required, and convert values into appropriate units when necessary.	$G = 6.67 \times 10^{-11}$ N m^2 kg^{-2} $m_1 = 90$ kg $m_2 = 75$ kg $r = 80$ cm $= 0.80$ m
Substitute the values into the equation.	$F_g = 6.67 \times 10^{-11} \times \frac{90 \times 75}{0.80^2}$
Solve the equation.	$F_g = 7.0 \times 10^{-7}$ N

Worked example: Try yourself 1.1.1

GRAVITATIONAL ATTRACTION BETWEEN SMALL OBJECTS

Two bowling balls are sitting next to each other on a shelf so that the centres of the balls are 60 cm apart. Ball 1 has a mass of 7.0 kg and ball 2 has a mass of 5.5 kg. Calculate the force of gravitational attraction between them.

GRAVITATIONAL ATTRACTION BETWEEN MASSIVE OBJECTS

Gravitational forces between everyday objects are so small (as seen in Worked example 1.1.1) that they are hard to detect without specialised equipment and can usually be considered to be negligible.

For the gravitational force to become significant, at least one of the objects must have a very large mass—for example, a planet (see Figure 1.1.4).

FIGURE 1.1.4 Gravitational forces become significant when at least one of the objects has a large mass, for example the Earth and the Moon.

Worked example 1.1.2

GRAVITATIONAL ATTRACTION BETWEEN MASSIVE OBJECTS

Calculate the force of gravitational attraction between the Sun and the Earth given the following data:

$m_{Sun} = 2.0 \times 10^{30}$ kg

$m_{Earth} = 6.0 \times 10^{24}$ kg

$r_{Sun-Earth} = 1.5 \times 10^{11}$ m

Thinking	Working
Recall the formula for Newton's law of universal gravitation.	$F_g = G\frac{m_1 m_2}{r^2}$
Identify the information required.	$G = 6.67 \times 10^{-11}$ N m^2 kg^{-2} $m_1 = 2.0 \times 10^{30}$ kg $m_2 = 6.0 \times 10^{24}$ kg $r = 1.5 \times 10^{11}$ m
Substitute the values into the equation.	$F_g = 6.67 \times 10^{-11} \times \frac{2.0 \times 10^{30} \times 6.0 \times 10^{24}}{(1.5 \times 10^{11})^2}$
Solve the equation.	$F_g = 3.6 \times 10^{22}$ N

Worked example: Try yourself 1.1.2

GRAVITATIONAL ATTRACTION BETWEEN MASSIVE OBJECTS

Calculate the force of gravitational attraction between the Earth and the Moon, given the following data:

$m_{Earth} = 6.0 \times 10^{24}$ kg

$m_{Moon} = 7.3 \times 10^{22}$ kg

$r_{Moon–Earth} = 3.8 \times 10^{8}$ m

The forces in Worked example 1.1.2 are much greater than those in Worked example 1.1.1, illustrating the difference in the gravitational force when at least one of the objects has a very large mass.

EXTENSION

Understanding the structure of the universe

In the century before Newton, there had been some controversy about the structure of the universe. In 1543, the commonly accepted geocentric (i.e. Earth-centred) model of the universe had been challenged by a Polish astronomer called Nicolaus Copernicus. He proposed that the Sun was the centre of the universe. Unfortunately, some faulty assumptions meant that the predictions of Copernicus' Sun-centred or heliocentric model (shown in Figure 1.1.5) did not match observations any better than the geocentric model.

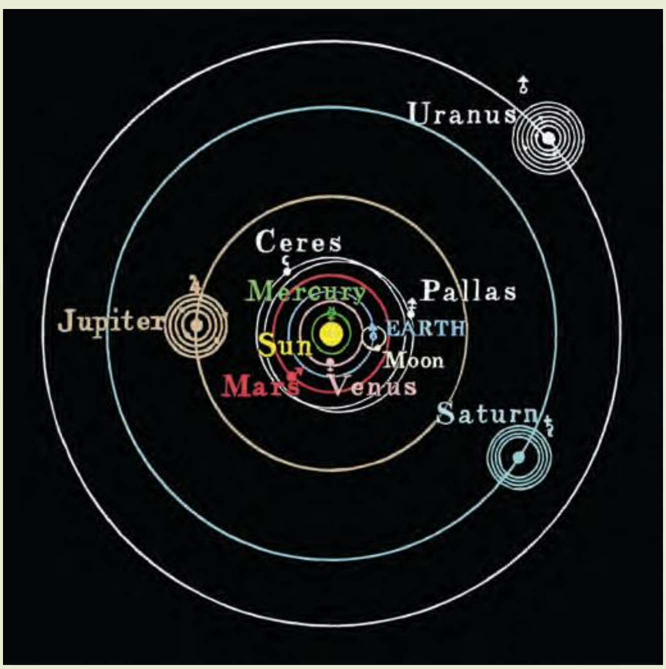


FIGURE 1.1.5 Nicolaus Copernicus' proposed heliocentric model of the solar system.

The Danish astronomer Tycho Brahe had been observing and studying the heavens for many years, accumulating a comprehensive collection of data. According to Brahe's documentation, his assistant, German mathematician Johannes Kepler, refined the Copernican model to reflect actual observations.

Through these calculations, Kepler discovered that the orbit of the planets around the Sun was elliptical and not circular as previously thought (see Figure 1.1.6). At the time, this discovery challenged conventional beliefs about the 'perfection' of heavenly bodies, and, as a consequence, Kepler's ideas were not widely accepted. In fact, in some countries his books were banned and publicly burned.

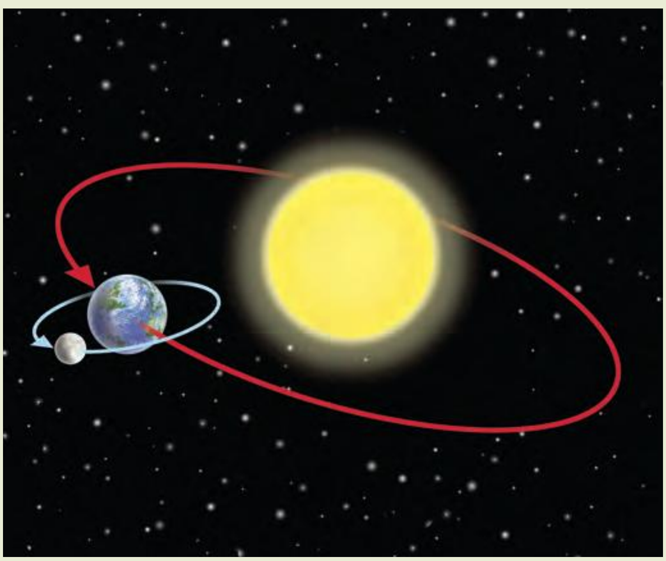


FIGURE 1.1.6 Johannes Kepler discovered that the orbit of planets around the Sun was elliptical.

One of Newton's great achievements was that he was able to use his law of universal gravitation to mathematically derive all of Kepler's planetary laws.

This allowed Newton to accurately explain the motion of the planets in terms of gravitational attraction. Within a few years of the publication of Newton's work, the geocentric model had largely been abandoned in favour of the heliocentric model.

FIGURE 1.1.7 The Earth and Moon exert gravitational forces on each other.

EFFECT OF GRAVITY

According to Newton's third law of motion, forces occur in action–reaction pairs. An example of such a pair is shown in Figure 1.1.7. The Earth exerts a gravitational force on the Moon and, conversely, the Moon exerts an equal and opposite force on the Earth. Using Newton's second law of motion, you can see that the effect of the gravitational force of the Moon on the Earth will be much smaller than the corresponding effect of the Earth on the Moon. This is because of the Earth's larger mass.

Worked example 1.1.3

ACCELERATION CAUSED BY A GRAVITATIONAL FORCE

The force of gravitational attraction between the Moon and the Earth is approximately 2.0×10^{20} N. Calculate the acceleration of the Earth and the Moon caused by this force. Compare these accelerations by calculating the ratio $\frac{a_{Moon}}{a_{Earth}}$.

Use the following data:

$m_{Earth} = 6.0 \times 10^{24}$ kg

$m_{Moon} = 7.3 \times 10^{22}$ kg

Thinking	Working
Recall the formula for Newton's second law of motion.	$F = ma$
Transpose the equation to make *a* the subject.	$a = \frac{F}{m}$
Substitute values into this equation to find the accelerations of the Moon and the Earth.	$a_{Earth} = \frac{2.0 \times 10^{20}}{6.0 \times 10^{24}} = 3.3 \times 10^{-5}$ m s^{-2} $a_{Moon} = \frac{2.0 \times 10^{20}}{7.3 \times 10^{22}} = 2.7 \times 10^{-3}$ m s^{-2}
Compare the two accelerations.	$\frac{a_{Moon}}{a_{Earth}} = \frac{2.7 \times 10^{-3}}{3.3 \times 10^{-5}} = 82$ The acceleration of the Moon is 82 times greater than the acceleration of the Earth.

Worked example: Try yourself 1.1.3

ACCELERATION CAUSED BY A GRAVITATIONAL FORCE

The force of gravitational attraction between the Sun and the Earth is approximately 3.6×10^{22} N. Calculate the acceleration of the Earth and the Sun caused by this force. Compare these accelerations by calculating the ratio $\frac{a_{Earth}}{a_{Sun}}$.

Use the following data:

$m_{Earth} = 6.0 \times 10^{24}$ kg

$m_{Sun} = 2.0 \times 10^{30}$ kg

Gravity in the solar system

Although the accelerations caused by gravitational forces in Worked example 1.1.3 are small, over billions of years they created the motion of the solar system.

In the Earth–Moon system, the acceleration of the Moon is many times greater than that of the Earth, which is why the Moon orbits the Earth. Although the Moon's gravitational force causes a much smaller acceleration of the Earth, it does have other significant effects, such as the tides.

Similarly, the Earth and other planets orbit the Sun because their masses are much smaller than the Sun's mass. However, the combined gravitational effect of the planets of the solar system (and Jupiter in particular) causes the Sun to wobble slightly as the planets orbit it.

PHYSICSFILE

Extrasolar planets

In recent years, scientists have been interested in discovering whether other stars have planets like those in our own solar system. One of the ways in which these 'extrasolar planets' (or 'exoplanets') can be detected is from their gravitational effect.

When a large planet (i.e. Jupiter-sized or larger) orbits a star, it causes the star to wobble. This causes variations in the star's appearance, which can be detected on Earth. Hundreds of exoplanets have been discovered using this technique.

WEIGHT AND GRAVITATIONAL FORCE

In Unit 2 Physics the **weight** of an object was calculated using the formula $W = F_g = mg$. Weight is another name for the gravitational force acting on an object near the Earth's surface.

Worked example 1.1.4 below shows that the formula $F_g = mg$ and Newton's law of universal gravitation give the same answer for the gravitational force acting on objects on the Earth's surface. It is important to note that the distance used in these calculations is the distance between the centres of the two objects, which is effectively the radius of the Earth.

Worked example 1.1.4

GRAVITATIONAL FORCE AND WEIGHT

Compare the weight of an 80 kg person calculated using $F_g = mg$ with the gravitational force calculated using $F_g = G\frac{m_1m_2}{r^2}$.

Use the following dimensions of the Earth in your calculations:

$g = 9.8\ \text{m s}^{-2}$

$m_{Earth} = 6.0 \times 10^{24}$ kg

$r_{Earth} = 6.4 \times 10^6$ m

Thinking	Working
Apply the weight equation.	$F_g = mg$ $= 80 \times 9.8$ $= 784$ N $= 780$ N (to two significant figures)
Apply Newton's law of universal gravitation.	$F_g = G\frac{m_1m_2}{r^2}$ $= 6.67 \times 10^{-11} \times \frac{6.0 \times 10^{24} \times 80}{(6.4 \times 10^6)^2}$ $= 780$ N
Compare the two values.	Both equations give the same result to two significant figures.

Worked example: Try yourself 1.1.4

GRAVITATIONAL FORCE AND WEIGHT

Compare the weight of a 1.0 kg mass on the Earth's surface calculated using the formulas $F_g = mg$ and $F_g = G\frac{m_1m_2}{r^2}$. Use the following dimensions of the Earth where necessary:

$g = 9.8\ \text{m s}^{-2}$

$m_{Earth} = 6.0 \times 10^{24}$ kg

$r_{Earth} = 6.4 \times 10^6$ m

Worked example 1.1.4 shows that the constant for the **acceleration due to gravity**, g, can be derived directly from the dimensions of the Earth. An object with mass m sitting on the surface of the Earth is a distance of 6.4×10^6 m from the centre of the Earth.

Given that the Earth has a mass of 6.0×10^{24} kg, then:

$$\text{Weight} = F_g$$

$$\therefore mg = G\frac{m_{Earth}m}{(r_{Earth})^2}$$

$$= mG\frac{m_{Earth}}{(r_{Earth})^2}$$

$$\therefore g = G\frac{m_{Earth}}{(r_{Earth})^2}$$

$$= 6.67 \times 10^{-11} \times \frac{6.0 \times 10^{24}}{(6.4 \times 10^6)^2}$$

$$= 9.8\ \text{m s}^{-2}$$

EXTENSION

Multi-body systems

So far, only gravitational systems involving two objects have been considered, such as the Moon and the Earth. In reality, objects experience gravitational force from every other object around them. Usually, most of these forces are negligible and only the gravitational effect of the largest object nearby (i.e. the Earth) needs to be considered.

When there is more than one significant gravitational force acting on a body, the gravitational forces must be added together as vectors to determine the net gravitational force (see Figure 1.1.8).

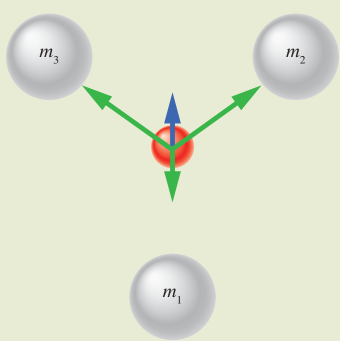


FIGURE 1.1.8 For the three masses $m_1 = m_2 = m_3$, the gravitational forces acting on the central red ball are shown by the green arrows. The vector sum of the green arrows is shown by the blue arrow. This will be the direction of the net (or resultant) gravitational force on the red ball due to the other three masses.

The direction and relative magnitude of the net gravitational force in a multi-body system depends entirely on the masses and positions of the attracting objects (i.e. m_1, m_2 and m_3 in Figure 1.1.8).

So, the rate of acceleration of objects near the surface of the Earth is a result of the Earth's mass and radius. A planet with a different mass and/or different radius will therefore have a different value for g. Likewise, if an object is above the Earth's surface, the value of r will be greater and the value of g will be smaller (due to the inverse square law). This is why the strength of the Earth's gravity reduces as you travel away from the Earth.

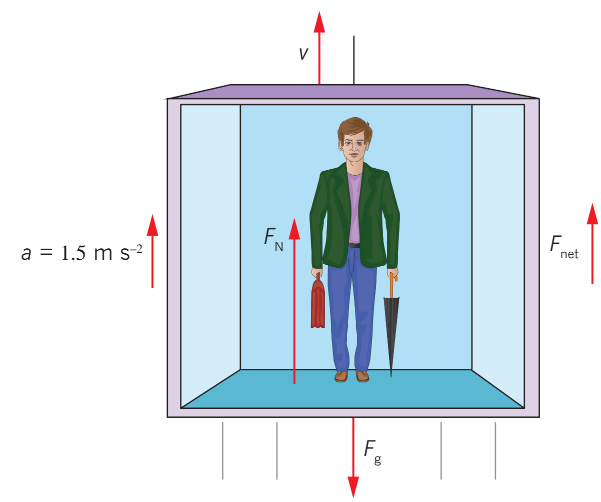
APPARENT WEIGHT

Scientists use the term 'weight' simply to mean 'the force due to gravity'. It is also correct to interpret weight as the contact force (or **normal reaction force**) between an object and the Earth's surface. In most situations these two definitions are effectively the same; however there are some cases, for example when a person is accelerating up or down in an elevator, where they give different results. In these situations, the normal force (F_N) is referred to as the **apparent weight** since this is the force that the person will feel through their feet.

Worked example 1.1.5

APPARENT WEIGHT

A 74 kg person is standing in an elevator which is accelerating upwards at 1.5 m s^{-2}. Calculate the weight and apparent weight of the person. Use g = 9.8 m s^{-2}.



Thinking	Working
Calculate the weight of the person using $F_g = mg$.	$F_g = mg = 74 \times 9.8 = 725$ N
Calculate the force required to accelerate the person upwards at 1.5 m s^{-2}.	$F_{net} = ma = 74 \times 1.5 = 111$ N
The net force that causes the acceleration results from the normal reaction force (upwards) and the weight force (downwards). Since the elevator is accelerating upwards, $F_N > F_g$. Recall that the normal reaction force gives the apparent weight.	$F_{net} = 111$ $F_N - F_g = 111$ $F_N - 725 = 111$ $F_N = 725 + 111$ F_N = apparent weight = 836 N

Worked example: Try yourself 1.1.5

APPARENT WEIGHT

Calculate the apparent weight of a 90 kg person in an elevator which is accelerating downwards at 0.8 m s^{-2}. Use g = 9.8 m s^{-2}.

1.1 Review

SUMMARY

- All objects with mass attract one another with a gravitational force.
- The gravitational force acts equally on each of the masses.
- The magnitude of the gravitational force is given by Newton's law of universal gravitation:
  $F_g = G\frac{m_1 m_2}{r^2}$
- Gravitational forces are usually negligible unless one of the objects is massive, e.g. a planet.
- The weight of an object on the Earth's surface is due to the gravitational attraction of the Earth, i.e. weight = F_g.
- The acceleration due to gravity of an object near the Earth's surface can be calculated using the dimensions of the Earth:
  $g = G\frac{m_{Earth}}{(r_{Earth})^2} = 9.8\ \text{m s}^{-2}$
- Objects can have an apparent weight that is greater or less than their normal weight. This occurs when they are accelerating vertically.

KEY QUESTIONS

1 What are the proportionalities in Newton's law of universal gravitation?

2 What does the symbol *r* represent in Newton's law of universal gravitation?

3 Calculate the force of gravitational attraction between the Sun and Mars given the following data:
$m_{Sun} = 2.0 \times 10^{30}$ kg
$m_{Mars} = 6.4 \times 10^{23}$ kg
$r_{Sun–Mars} = 2.2 \times 10^{11}$ m

4 The force of gravitational attraction between the Sun and Mars is 1.8×10^{21} N. Calculate the acceleration of Mars given that $m_{Mars} = 6.4 \times 10^{23}$ kg.

5 On 14 April 2014, Mars came within 93 million km of Earth. Its gravitational effect on the Earth was the strongest it had been for over 6 years. Use the following data to answer the questions below.
$m_{Sun} = 2.0 \times 10^{30}$ kg
$m_{Earth} = 6.0 \times 10^{24}$ kg
$m_{Mars} = 6.4 \times 10^{23}$ kg

a Calculate the gravitational force between the Earth and Mars on 14 April 2014.

b Calculate the force of the Sun on the Earth if the distance between them was 153 million km.

c Compare your answers to parts (a) and (b) above by expressing the Mars–Earth force as a percentage of the Sun–Earth force.

6 The acceleration of the Moon caused by the gravitational force of the Earth is much larger than the acceleration of the Earth due to the gravitational force of the Moon. What is the reason for this?

7 Calculate the acceleration of an object dropped near the surface of Mercury if this planet has a mass of 3.3×10^{23} kg and a radius of 2500 km. Assume that the gravitational acceleration on Mercury can be calculated similarly to that on Earth.

8 Calculate the weight of a 65 kg cosmonaut standing on the surface of Mars, given that the planet has a mass of 6.4×10^{23} kg and a radius of 3.4×10^6 m.

9 In your own words, explain the difference between the terms weight and apparent weight, giving an example of a situation where the magnitudes of these two forces would be different.

10 Calculate the apparent weight of a 50 kg person in an elevator under the following circumstances.

a accelerating upwards at 1.2 m s^{-2}

b moving upwards at a constant speed of 5 m s^{-1}

1.2 Gravitational fields

Newton's law of universal gravitation describes the force acting between two mutually attracting bodies. In reality, complex systems like the solar system involve a number of objects (i.e. the Sun and planets shown in Figure 1.2.1) that are all exerting attractive forces on each other at the same time.

In the 18th century, to simplify the process of calculating the effect of simultaneous gravitational forces, scientists developed a mental construct known as the gravitational field. In the following centuries, the idea of a **field** was also applied to other forces and has become a very important concept in physics.

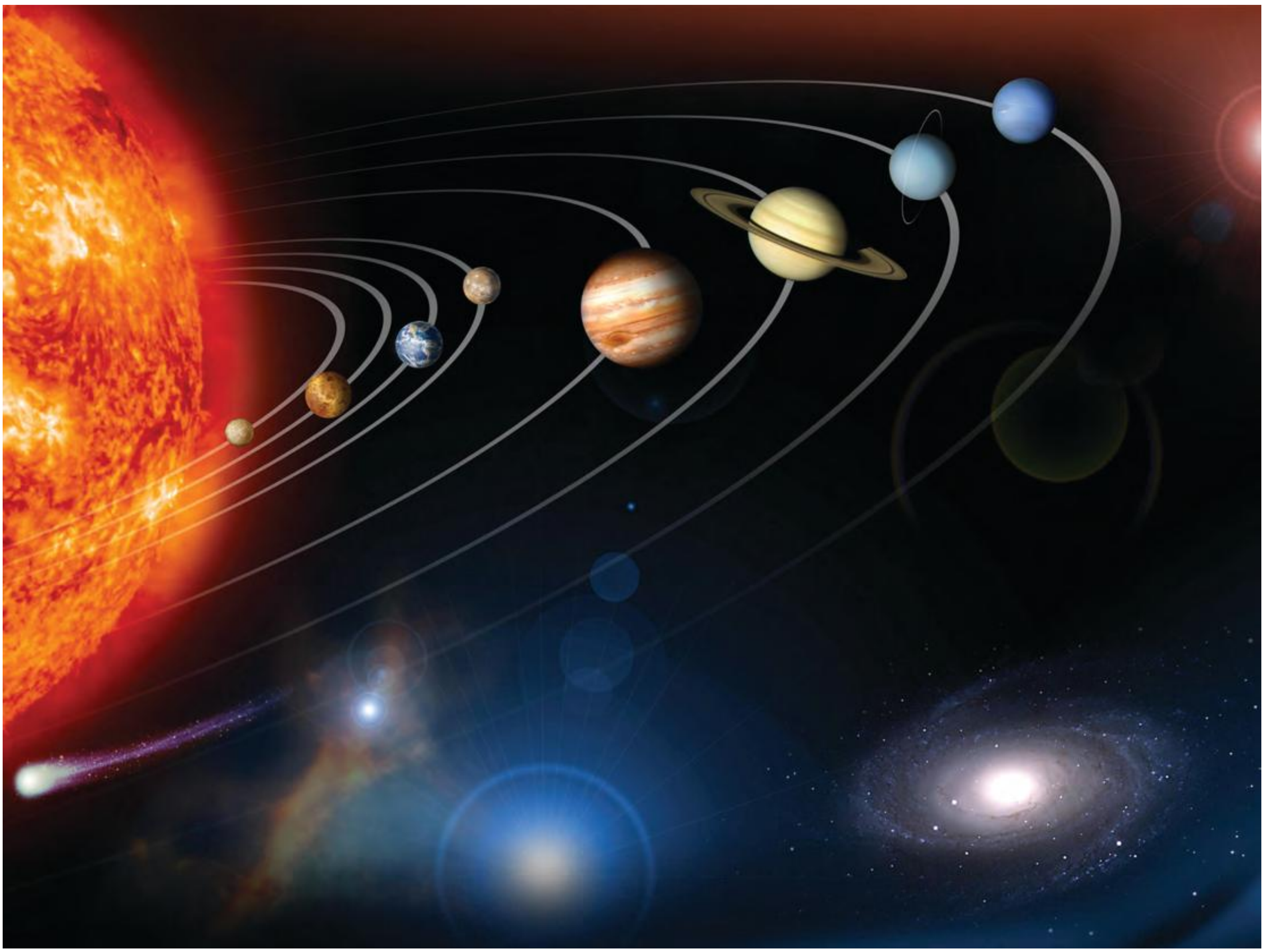


FIGURE 1.2.1 The solar system is a complex gravitational system.

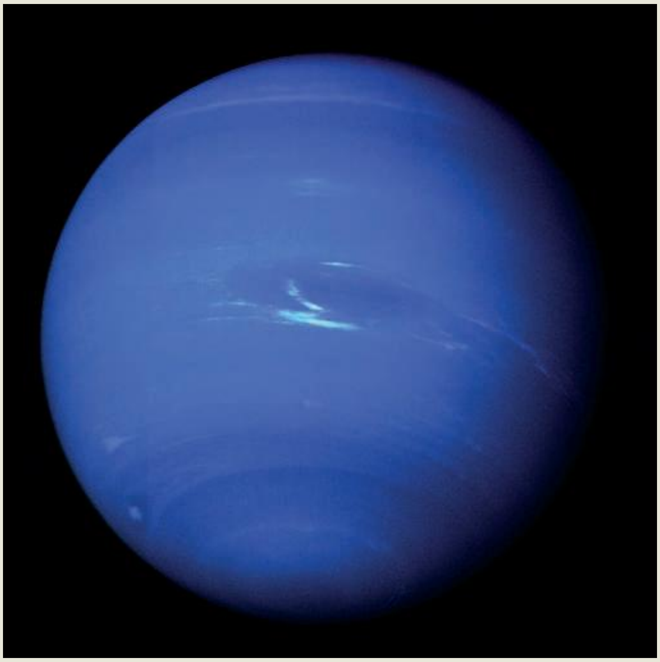
GRAVITATIONAL FIELDS

A **gravitational field** is a region in which a gravitational force is exerted on all matter within that region. Every physical object has an accompanying gravitational field. For example, the space around your body contains a gravitational field because any other object that comes into this region will experience a (small) force of gravitational attraction to your body.

The gravitational field around a large object like a planet is much more significant than that around a small object. The Earth's gravitational field exerts a significant influence on objects on its surface and even up to thousands of kilometres into space.

PHYSICS IN ACTION

Discovery of Neptune



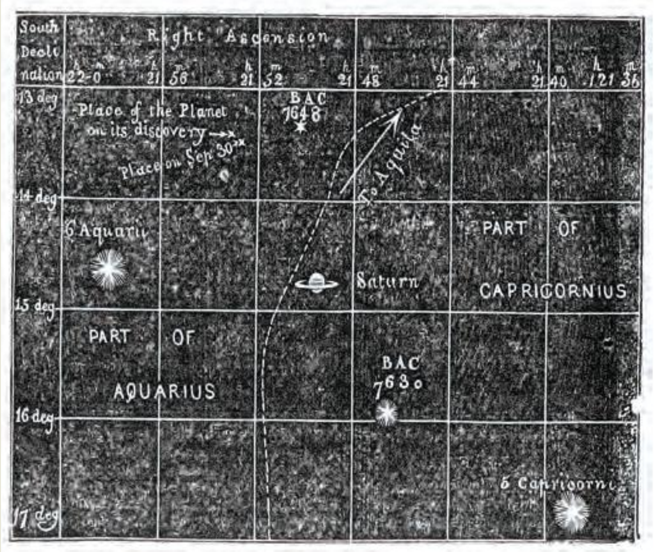


FIGURE 1.2.2 This star chart published in 1846 shows the location of Neptune in the constellation Aquarius when it was discovered on 23 September, and its location one week later.

The planet Neptune was discovered through its gravitational effect on other planets. Two astronomers, Urbain Le Verrier of France and John Couch Adams of England, each independently identified that the observed orbit of Uranus varied significantly from predictions made based on the gravitational effects of the Sun and other known planets. Both suggested that this was due to the influence of a distant, undiscovered planet.

Le Verrier sent a prediction of the location of the new planet to Gottfried Galle at the Berlin Observatory and, on 23 September 1846, Neptune was discovered within 1° of Le Verrier's prediction (see Figure 1.2.2).

Representing gravitational fields

Over time, scientists have developed a commonly understood method of representing fields using a series of arrows known as field lines (see Figure 1.2.3). For gravitational fields, these are constructed as follows:

- the direction of the arrowhead indicates the direction of the gravitational force
- the space between the arrows indicates the relative magnitude of the field:
  - closely spaced arrows indicate a strong field
  - widely spaced arrows indicate a weaker field
  - parallel field lines indicate constant or **uniform** field strength.

An infinite number of field lines could be drawn, so only a few are chosen to represent the rest. The size of the gravitational force acting on a mass in the region of a gravitational field is determined by the strength of the field, and the force acts in the direction of the field.

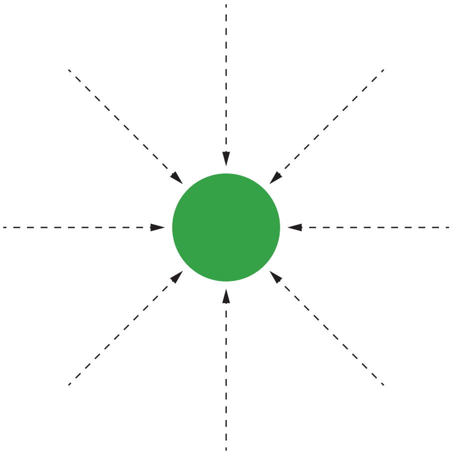
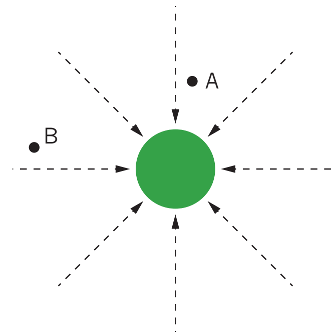


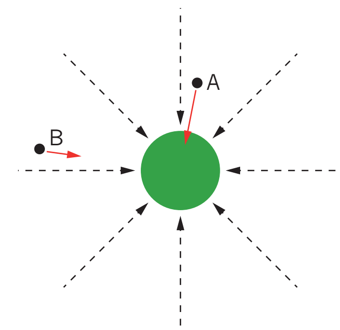
FIGURE 1.2.3 The arrows in this gravitational field diagram indicate that objects will be attracted towards the mass in the centre; the spacing of the lines shows that force will be strongest at the surface of the central mass and weaker further away from it.

Worked example 1.2.1

INTERPRETING GRAVITATIONAL FIELD DIAGRAMS

The diagram below shows the gravitational field of a moon.



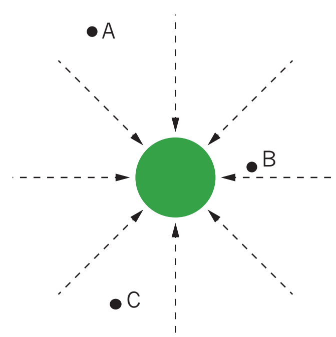
a Use arrows to indicate the direction of the gravitational force acting at points A and B.	
Thinking	**Working**
The direction of the field arrows indicates the direction of the gravitational force, which is inwards towards the centre of the moon.	

b Indicate the relative strength of the gravitational field at each point.	
Thinking	**Working**
The closer the field lines, the stronger the force. The field lines are closer together at point A than they are at point B, as point A is closer to the moon.	The field is stronger at point A than at point B.

Worked example: Try yourself 1.2.1

INTERPRETING GRAVITATIONAL FIELD DIAGRAMS

The diagram below shows the gravitational field of a planet.



a Use arrows to indicate the direction of the gravitational force acting at points A, B and C.

b Indicate the relative strength of the gravitational field at each point.

GRAVITATIONAL FIELD STRENGTH

In theory, gravitational fields extend infinitely out into space. However, since the magnitude of the gravitational force decreases with the square of distance, eventually these fields become so weak as to become negligible.

In Section 1.1, it was shown that it is possible to calculate the acceleration due to gravity of objects near the Earth's surface using the dimensions of the Earth:

$$g = G\frac{m_{\text{Earth}}}{(r_{\text{Earth}})^2} = 9.8 \text{ m s}^{-2}$$

The constant g can also be used as a measure of the strength of the gravitational field. When understood in this way, the constant is written with the equivalent units of N kg^{-1} rather than m s^{-2}. This means $g_{\text{Earth}} = 9.8$ N kg^{-1}.

These units indicate that objects on the surface of the Earth experience 9.8 N of gravitational force for every kilogram of their mass.

Accordingly, the familiar equation $F_g = mg$ can be transposed so that the **gravitational field strength**, g, can be calculated:

$$g = \frac{F_g}{m}$$

where g is gravitational field strength (N kg^{-1})

F_g is the force due to gravity (N)

m is the mass of an object in the field (kg)

PHYSICS FILE

N kg^{-1} = m s^{-2}

It is a simple matter to show that N kg^{-1} and m s^{-2} are equivalent units.

From Newton's second law, $F = ma$, you will remember that:

1 N = 1 kg m s^{-2}

$\therefore$ 1 N kg^{-1} = 1 kg m s^{-2} × kg^{-1}

= 1 m s^{-2}

Worked example 1.2.2

CALCULATING GRAVITATIONAL FIELD STRENGTH

When a student hangs a 1 kg mass from a spring balance, the balance measures a downwards force of 9.8 N.

According to this experiment, what is the gravitational field strength of the Earth in this location?

Thinking	Working
Recall the equation for gravitational field strength.	$g = \frac{F_g}{m}$
Substitute in the appropriate values.	$g = \frac{9.8}{1}$
Solve the equation.	$g = 9.8$ N kg^{-1}

Worked example: Try yourself 1.2.2

CALCULATING GRAVITATIONAL FIELD STRENGTH

A student uses a spring balance to measure the weight of a piece of wood as 2.5 N.

If the piece of wood is thought to have a mass of 260 g, calculate the gravitational field strength indicated by this experiment.

The formula for gravitational field strength, $g = \frac{F_g}{m}$, can be combined with Newton's law of universal gravitation, $F_g = G\frac{Mm}{r^2}$, to develop the formula for gravitational field strength:

$$g = \frac{F_g}{m} = \frac{\left(G\frac{Mm}{r^2}\right)}{m}$$

Therefore:

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

where g is the gravitational field strength (N kg^{-1})

G is the gravitational constant, 6.67×10^{-11} N m^2 kg^{-2}

M is the mass of the planet or moon (the central body; kg)

r is the radius of the planet or moon (m)

Inverse square law

The concept of a field is a very powerful tool for understanding forces that act at a distance. It has also been applied to forces such as the electrostatic force between charged objects and the force between two magnets.

The study of gravitational fields introduces the concept of the inverse square law. From the point source of a field, whether it be gravitational, electric or magnetic, the field will spread out radially in three dimensions. When the distance from the source is doubled, the field will be spread over four times the original area.

In Figure 1.2.4, going from r to $2r$ to $3r$, the area shown increases from one square to four squares (2^2) to nine squares (3^2). Using the inverse part of the inverse square law, at a distance $2r$ the strength of the field will be reduced to a quarter of that at r, as is the force that the field would exert. At $3r$ from the source, the field will be reduced to one-ninth of that at the source, and so on.

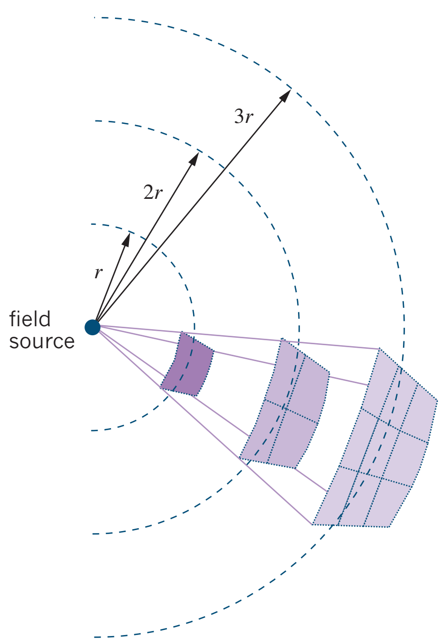


FIGURE 1.2.4 As the distance from the source of a field increases, the field is spread over an area that increases with the square of the distance from the source, resulting in the strength of the field decreasing by the same ratio.

In terms of the gravitational field, the strength of the force varies inversely with the distance between the objects squared:

$$F \propto \frac{M}{r^2}$$

where F is the force and r is the distance from the source of the gravitational field.

This is referred to as the inverse square law.

One key difference between the gravitational force and other inverse square forces is that the gravitational force is always attractive, whereas like charges or magnets repel one another.

Inverse square laws are an important concept in physics, not only in the study of fields but also for other phenomena where energy is moving away from its source in three dimensions, such as in sound and other waves.

Variations in gravitational field strength of the Earth

The gravitational field strength of the Earth, g, is usually assigned a value of 9.81 N kg^{-1}. However, the field strength experienced by objects on the surface of the Earth can actually vary between 9.76 N kg^{-1} and 9.83 N kg^{-1}, depending on the location.

PHYSICSFILE

Variations in gravitational field strength

The Earth's gravitational field strength is not the same at every point on the Earth's surface. As the Earth is not a perfect sphere, objects near the equator are slightly further from the centre of the Earth than objects at the poles. This means that the Earth's gravitational field is slightly stronger at the poles than at the equator.

Geological formations can also create differences in gravitational field strength, depending on their composition. Geologists use a sensitive instrument known as a **gravimeter** (see Figure 1.2.6) that detects small local variations in gravitational field strength to indicate underground features. Rocks with above-average density, such as those containing mineral ores, create slightly stronger gravitational fields, whereas less-dense sedimentary rocks produce weaker fields.

FIGURE 1.2.6 A gravimeter can be used to measure the strength of the local gravitational field.

If the surface of the Earth is considered a flat surface as it appears in everyday life, then the gravitational field lines are approximately parallel, indicating a uniform field (see Figure 1.2.7).

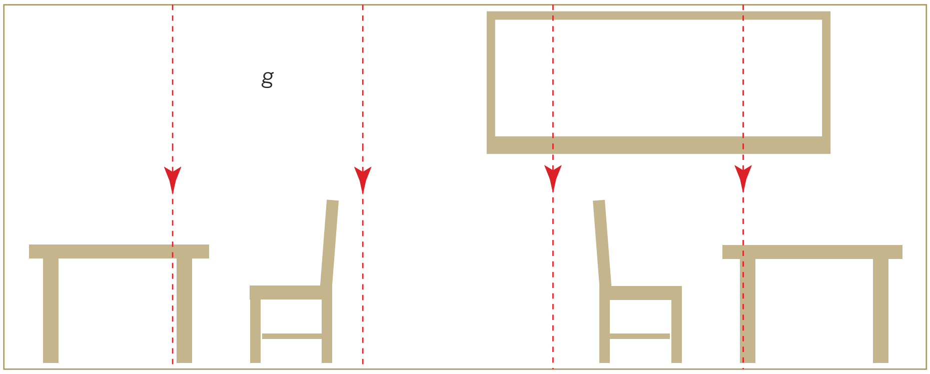


FIGURE 1.2.7 The uniform gravitational field, g, is represented by evenly spaced parallel lines in the direction of the force.

PHYSICSFILE

The shape of the Earth

The shape of the Earth is known as an oblate spheroid (see Figure 1.2.5). Mathematically, this is the shape that's made when an ellipse is rotated around its minor axis. The diameter of the Earth between the North and South poles is approximately 40 km shorter than its diameter at the equator.

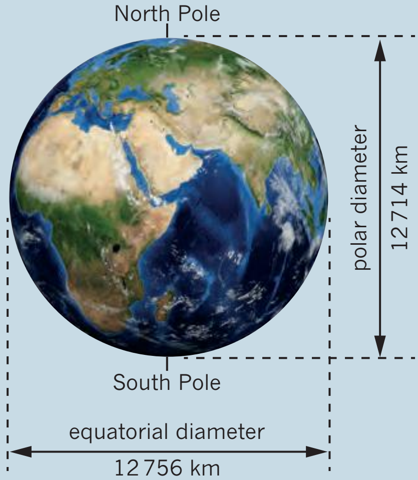


FIGURE 1.2.5 The Earth is a flattened sphere, which means its gravitational field is slightly stronger at the poles.

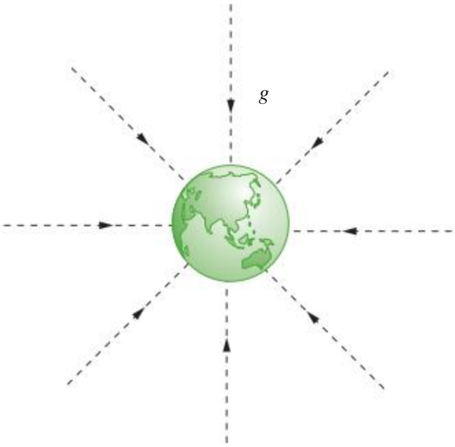


FIGURE 1.2.8 The Earth's gravitational field becomes progressively weaker out into space.

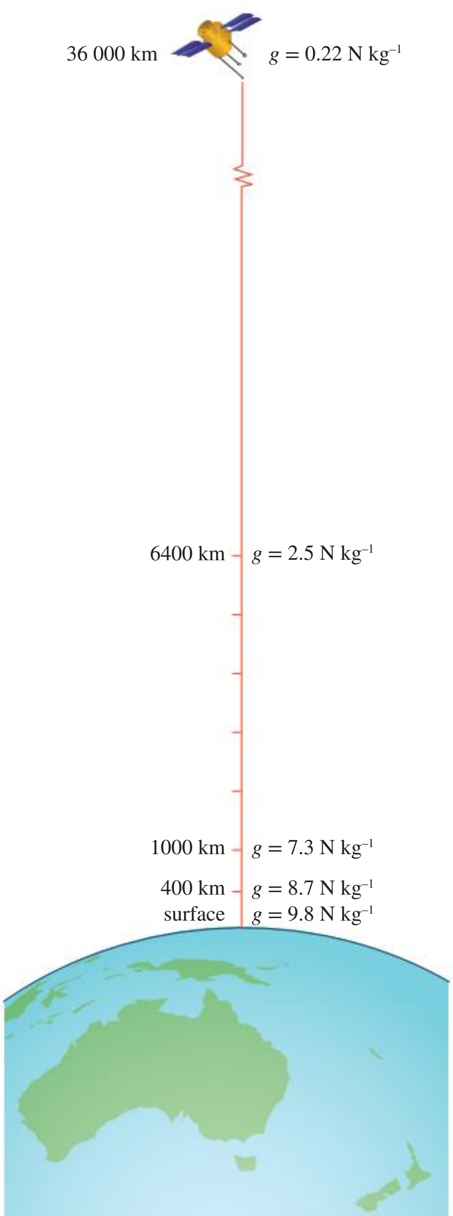


FIGURE 1.2.9 The Earth's gravitational field strength is weaker at higher altitudes.

However, when the Earth is viewed from a distance as a sphere, it becomes clear that the Earth's gravitational field is not uniform at all (see Figure 1.2.8). The increasing distance between the field lines indicates that the field becomes progressively weaker out into space.

This is because gravitational field strength, like gravitational force, is governed by the inverse square law:

$$g = G\frac{M_{\text{Earth}}}{(r_{\text{Earth}})^2}$$

The gravitational field strength at different altitudes can be calculated by adding the **altitude** to the radius of the Earth to calculate the distance of the object from Earth's centre (see Figures 1.2.9 and 1.2.10).

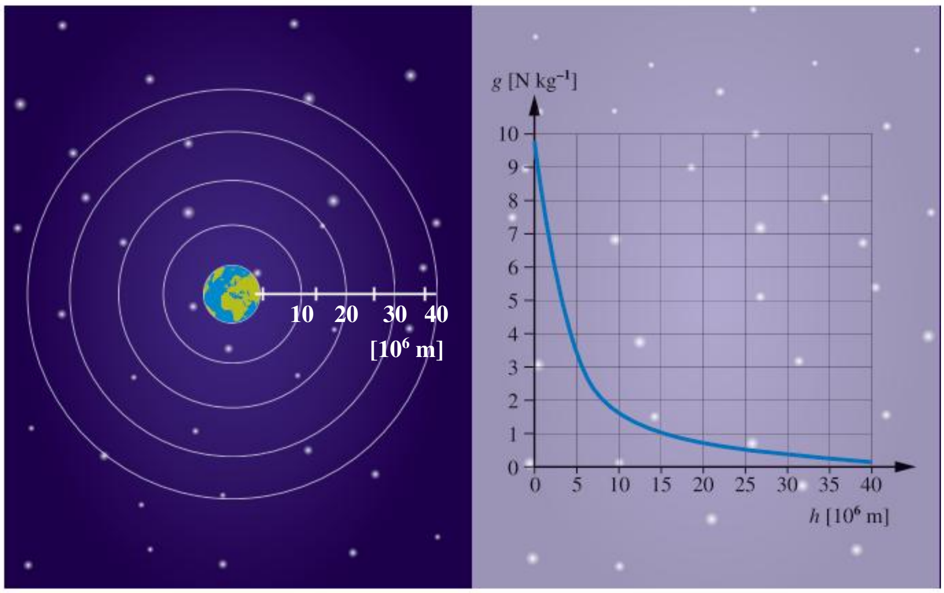


FIGURE 1.2.10 As the distance from the surface of the Earth is increased from 0 to 40 × 10^6 m, the value for *g* decreases rapidly from 9.8 N kg^{-1}, according to the inverse square law. The blue line on the graph gives the value of *g* at various altitudes (*h*).

$$g = \frac{GM_{\text{Earth}}}{(r_{\text{Earth}} + \text{altitude})^2}$$

Worked example 1.2.3

CALCULATING GRAVITATIONAL FIELD STRENGTH AT DIFFERENT ALTITUDES

Calculate the strength of the Earth's gravitational field at the top of Mt Everest using the following data:

$r_{\text{Earth}} = 6.38 \times 10^6$ m

$m_{\text{Earth}} = 5.97 \times 10^{24}$ kg

height of Mt Everest = 8850 m

Thinking	Working
Recall the formula for gravitational field strength.	$g = G\frac{M}{r^2}$
Add the height of Mt Everest to the radius of the Earth.	$r = 6.38 \times 10^6 + 8850$ m $= 6.389 \times 10^6$ m
Substitute the values into the formula.	$g = G\frac{M}{r^2}$ $= 6.67 \times 10^{-11} \times \frac{5.97 \times 10^{24}}{(6.389 \times 10^6)^2}$ $= 9.76$ N kg^{-1}

Worked example: Try yourself 1.2.3

CALCULATING GRAVITATIONAL FIELD STRENGTH AT DIFFERENT ALTITUDES

Commercial airlines typically fly at an altitude of 11 000 m. Calculate the gravitational field strength of the Earth at this height using the following data:

$r_{Earth} = 6.38 \times 10^6$ m

$m_{Earth} = 5.97 \times 10^{24}$ kg

FIGURE 1.2.11 The gravitational field strength on the surface of Mars (shown here) is different to the gravitational field strength on the surface of the Earth, which, in turn, is different from that on the Moon.

Gravitational field strengths of other planets

The gravitational field strength on the surface of the Moon is much less than on Earth, at approximately 1.6 N kg^{-1}. This is because the Moon's mass is smaller than the Earth's.

The formula $g = G\frac{M}{r^2}$ can be used to calculate the gravitational field strength on the surface of any astronomical object, such as Mars (see Figure 1.2.11).

Worked example 1.2.4

GRAVITATIONAL FIELD STRENGTH ON ANOTHER PLANET OR MOON

Calculate the strength of the gravitational field on the surface of the Moon given that the Moon's mass is 7.35×10^{22} kg and its radius is 1740 km.
Give your answer correct to two significant figures.

Thinking	Working
Recall the formula for gravitational field strength.	$g = G\frac{M}{r^2}$
Convert the Moon's radius to m.	$r = 1740$ km $= 1740 \times 1000$ m $= 1.74 \times 10^6$ m
Substitute values into the formula.	$g = G\frac{M}{r^2}$ $= 6.67 \times 10^{-11} \times \frac{7.35 \times 10^{22}}{(1.74 \times 10^6)^2}$ $= 1.6$ N kg^{-1}

Worked example: Try yourself 1.2.4

GRAVITATIONAL FIELD STRENGTH ON ANOTHER PLANET OR MOON

Calculate the strength of the gravitational field on the surface of Mars.

$m_{Mars} = 6.42 \times 10^{23}$ kg

$r_{Mars} = 3390$ km

Give your answer correct to two significant figures.

PHYSICSFILE

Moon walking

Walking is a process of repeatedly stopping yourself from falling over. When astronauts first tried to walk on the Moon, they found that they fell too slowly to walk easily. Instead, they invented a kind of shuffling jump that was a much quicker way of moving around in the Moon's weak gravitational field. This type of 'moon walk' should not be confused with the famous dance move of the same name!

FIGURE 1.2.12 Astronauts had to invent a new way of walking to deal with the Moon's weak gravitational field.

1.2 Review

SUMMARY

- A gravitational field is a region in which a gravitational force is exerted on all matter within that region.
- A gravitational field can be represented by a gravitational field diagram:
  - The arrowheads indicate the direction of the gravitational force.
  - The spacing of the lines indicates the relative strength of the field. The closer the line spacing, the stronger the field.
- The strength of a gravitational field can be calculated using the following formulas:
  $g = \frac{F_g}{m}$ or $g = G\frac{M}{r^2}$
  The gravitational field strength on the Earth's surface is approximately 9.8 N kg^{-1}. This varies from location to location and with altitude.
- The gravitational field strength on the surface of any other planet depends on the mass and radius of the planet.

KEY QUESTIONS

1 Give the most appropriate unit for measuring gravitational field strength.

2 A group of students use a spring balance to measure the weight of a 150 g set of slotted masses to be 1.4 N. According to this measurement, what is the gravitational field strength in their classroom?

3 A gravitational field, g, is measured as 5.5 N kg^{-1} at a distance of 400 km from the centre of a planet. The distance from the centre of the planet is then increased to 1200 km. What would the ratio of the magnitude of the gravitational field be at this new distance compared to the original measurement?

4 Different types of satellite have different types of orbit, as shown in the table below. Calculate the strength of the Earth's gravitational field in each orbit.

r_{Earth} = 6380 km

m_{Earth} = 5.97 × 10^{24} kg

	Type of orbit	Altitude (km)
a	low Earth orbit	2000
b	medium Earth orbit	10000
c	semi-synchronous orbit	20200
d	geosynchronous orbit	35786

5 On 12 November 2014, the Rosetta spacecraft landed a probe on the comet 67P/Churyumov–Gerasimenko. Assuming this comet is a roughly spherical object with a mass of 1 × 10^{13} kg and a diameter of 1.8 km, calculate the gravitational field strength on its surface.

6 The masses and radii of three planets are given in the following table.

Planet	Mass (kg)	Radius (m)
Mercury	3.30 × 10^{23}	2.44 × 10^{6}
Saturn	5.69 × 10^{26}	6.03 × 10^{7}
Jupiter	1.90 × 10^{27}	7.15 × 10^{7}

Calculate the gravitational field strength, g, at the surface of each planet.

7 There are bodies outside our solar system, such as neutron stars, that produce very large gravitational fields. A typical neutron star can have a mass of 3.0 × 10^{30} kg and a radius of just 10 km. Calculate the gravitational field strength at the surface of such a star.

8 A newly discovered solid planet located in a distant solar system is found to be distinctly non-spherical in shape. Its polar radius is 5000 km, and its equatorial radius is 6000 km.
The gravitational field strength at the poles is 8.0 N kg^{-1}. How would the gravitational field strength at the poles compare with the strength at the equator?

9 There is a point between the Earth and the Moon where the total gravitational field is zero. The significance of this is that returning lunar missions are able to return to Earth under the influence of the Earth's gravitational field once they pass this point. Given that the mass of Earth is 6.0 ×10^{24} kg, the mass of the Moon is 7.3 ×10^{22} kg and the radius of the Moon's orbit is 3.8 ×10^{8} m, calculate the distance of this point from the centre of the Earth.

10 An astronaut travels away from Earth to a region in space where the gravitational force due to Earth is only 1.0% of that at Earth's surface. What distance, in Earth radii, is the astronaut from the centre of the Earth?

1.3 Work in a gravitational field

The concept of gravitational potential energy should be familiar to you from Unit 2 Physics. However, the nature of a gravitational field means that a more sophisticated understanding of gravitational potential energy is needed when considering the motion of objects like rockets or satellites (see Figure 1.3.1).

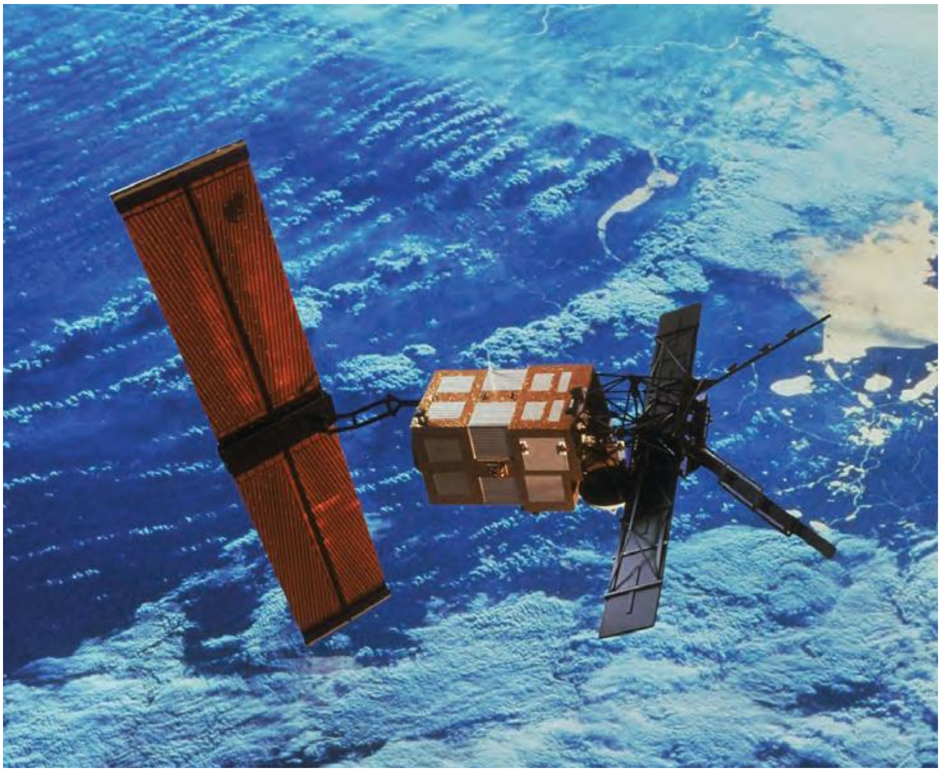


FIGURE 1.3.1 Satellites in orbit have gravitational potential energy.

REVISITING WORK AND CONSERVATION OF ENERGY

The **gravitational potential energy** of an object, E_g, is directly proportional to the mass of the object, m, its height above the surface of the planet, Δh, and the strength of the gravitational field, g. So:

$$E_g = mg\Delta h$$

where E_g is the gravitational potential energy of an object (J)

m is the mass of the object (kg)

g is the gravitational field strength (N kg^{-1}; 9.8 N kg^{-1} near the surface of the Earth)

Δh is the height of the object above a reference point (m)

The formula for gravitational potential energy is developed from the work–energy theorem, which assumes that work done against the force of gravity is converted into potential energy:

$$\Delta E = W = Fs$$

where ΔE is the change in gravitational potential energy (J)

W is the work done (J)

F is the force of gravity (N)

s is the distance moved in the gravitational field (m)

Worked example 1.3.1

WORK DONE FOR A CHANGE IN GRAVITATIONAL POTENTIAL ENERGY

A mountaineer climbs from a height of 700 m above sea level to the top of Mount Everest, which is 8848 m above sea level.

The total mass of the mountaineer (with equipment) is 100 kg. Assuming that the gravitational field strength of the Earth (g) is 9.8 N kg^{-1}, calculate the amount of work done (in MJ) by the mountaineer in climbing to the summit of the mountain.

Thinking	Working
Calculate the change in height.	$\Delta h = 8848 - 700$ $= 8148$ m
Substitute appropriate values into $E_g = mg\Delta h$. Remember to give your answer in MJ to two significant figures.	$E_g = mg\Delta h$ $= 100 \times 9.8 \times 8148$ $= 7\,985\,040$ J $= 8.0$ MJ
The work done by the mountaineer is equal to the change in gravitational potential energy.	$W = \Delta E = 8.0$ MJ

Worked example: Try yourself 1.3.1

WORK DONE FOR A CHANGE IN GRAVITATIONAL POTENTIAL ENERGY

Calculate the work done (in MJ) to lift a weather satellite of 3.2 tonnes from the Earth's surface to the limit of the atmosphere, which ends at the Karman line (exactly 100 km up from the surface of the Earth). Assume g = 9.8 N kg^{-1}.

Interplay between gravitational, kinetic and mechanical energy

Gravitational potential energy calculations are important because, when combined with the concepts of kinetic energy and conservation of mechanical energy, they allow the speed of a falling object to be determined.

Accordingly, we can define kinetic energy by the following equation:

$$E_k = \frac{1}{2}mv^2$$

where E_k is the kinetic energy of an object (J)

m is the mass of the object (kg)

v is the speed of the object (m s^{-1})

Worked example 1.3.2

SPEED OF A FALLING OBJECT

In a unique demonstration of Galileo's famous experiment, Apollo 15 astronaut Dave Scott simultaneously dropped a hammer and a feather while standing on the surface of the Moon (see Figure 1.3.2).

FIGURE 1.3.2 Astronaut Dave Scott dropping a hammer and feather on the Moon.

If the gravitational field strength on the Moon is 1.6 N kg^{-1}, the hammer had a mass of 450 g and it was dropped from a height of 1.4 m, calculate the speed of the hammer as it hit the Moon's surface.

Thinking	Working
Calculate the gravitational potential energy of the hammer on the Moon. Change the units of measurement when necessary.	$E_g = mg\Delta h$ $= 0.45 \times 1.6 \times 1.4$ $= 1.0$ J
Assume that when the hammer hit the surface of the Moon, all of its gravitational potential energy had been converted into kinetic energy.	$E_k = E_g = 1.0$ J
Use the definition of kinetic energy to calculate the speed of the hammer as it hit the ground.	$E_k = \frac{1}{2}mv^2$ $1.0 = \frac{1}{2} \times 0.45 \times v^2$ $\frac{1.0 \times 2}{0.45} = v^2$ $v = 2.1$ m s^{-1}

Worked example: Try yourself 1.3.2

SPEED OF A FALLING OBJECT

Calculate how fast a 450 g hammer would be going as it hit the ground if it were dropped from a height of 1.4 m on Earth, where $g = 9.8$ N kg^{-1}.

Work in a non-constant gravitational field

The formula $E_g = mg\Delta h$ is developed assuming that work is done against a constant force of gravity: $\Delta E = W = Fs$. While this assumption holds true for objects close to the surface of a planet, it is not adequate for objects like satellites that move to altitudes at which the gravitational field of the planet becomes significantly diminished.

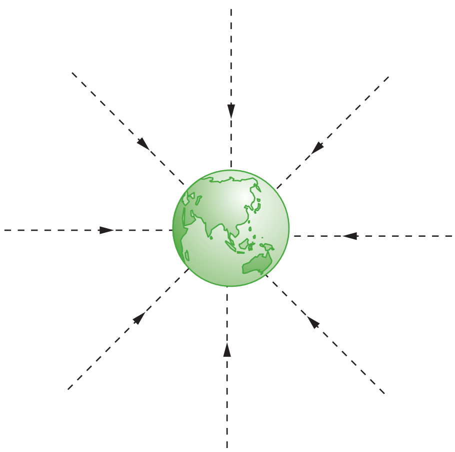


FIGURE 1.3.3 The Earth's gravitational field extends out into space, and the field is strongest close to the Earth where the lines are closer together.

Newton's law of universal gravitation indicates that the strength of the Earth's gravitational field changes with altitude: the field is stronger close to the ground and weaker at high altitudes (see Figure 1.3.3).

Consider the example of a 10 kg meteor falling towards the Earth from deep space as shown in Figure 1.3.4. Closer to the Earth, the meteor moves through regions of increasing gravitational field strength. So the gravitational force, F_g, on the meteor increases as it approaches Earth. Since the force is not constant, this means that the work done on the meteor (which corresponds to its change in gravitational potential energy) cannot be found by simply multiplying the gravitational force by the distance travelled.

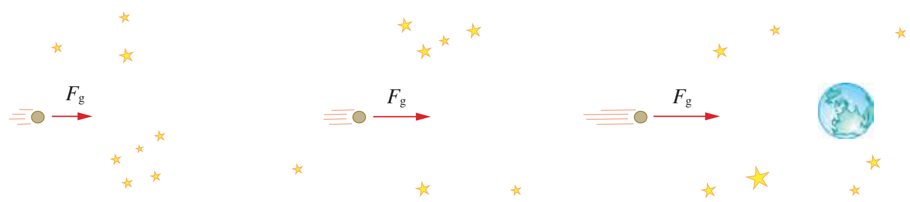


FIGURE 1.3.4 As a meteor approaches Earth, it moves through an increasingly stronger gravitational field and so is acted upon by a greater gravitational force.

Using the force–distance graph

When a free-falling body, like the meteor in Figure 1.3.4, is acted upon by a varying gravitational force, the energy changes of the body can be analysed by using a gravitational force–distance graph. As with other force–distance graphs, the area under the graph is equal to the work done, i.e. the energy change of the body. The area under the graph has units of newton metres (N m), which are equivalent to joules (J).

> **i** The area under a gravitational force–distance graph gives the change in energy that an object will experience as it moves through the gravitational field.

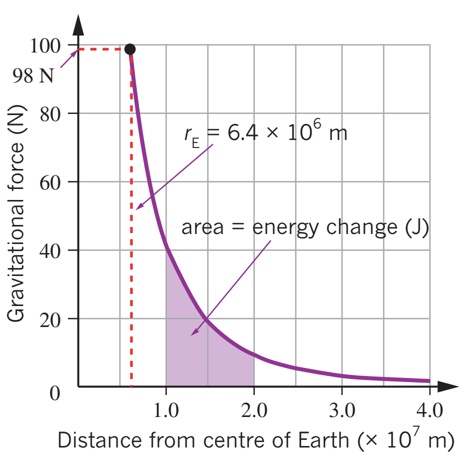


FIGURE 1.3.5 The gravitational force acting on a 10 kg meteor at different distances from the Earth. The shaded region represents the work done by the gravitational field as the body moves between 2.0×10^7 m and 1.0×10^7 m from the centre of the Earth.

The shaded area in Figure 1.3.5 represents the decrease in gravitational potential energy of the 10 kg meteor as it falls from a distance of 2.0×10^7 m to 1.0×10^7 m from the centre of the Earth. This area also represents the amount of kinetic energy that the meteor gains as it approaches Earth.

Note that the energy change of the meteor will be the same regardless of whether the meteor falls directly towards the planet (Figure 1.3.6(a)) or follows a more indirect path (Figure 1.3.6(b)).

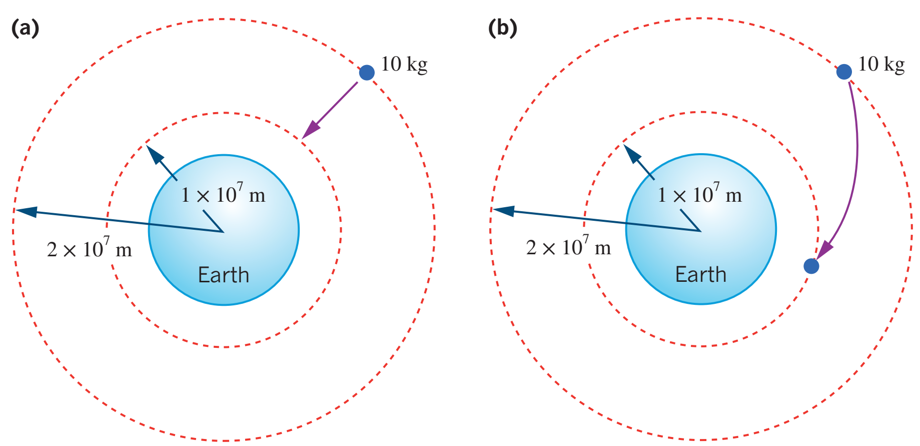


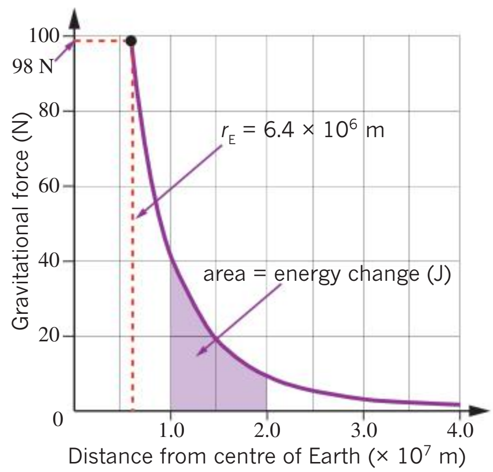
FIGURE 1.3.6 The shaded region on the gravitational force–distance graph in Figure 1.3.5 could represent the change in energy in the free-fall situations in either (a) or (b).

Worked example 1.3.3 shows how a force–distance graph can be used to determine the change in gravitational potential energy of a meteor.

Worked example 1.3.3

CHANGE IN GRAVITATIONAL POTENTIAL ENERGY USING A FORCE–DISTANCE GRAPH

A 10 kg meteor falls from a distance of 2.0×10^7 m to 1.0×10^7 m from the centre of the Earth. Use the graph below to determine the change in gravitational potential energy of the meteor.

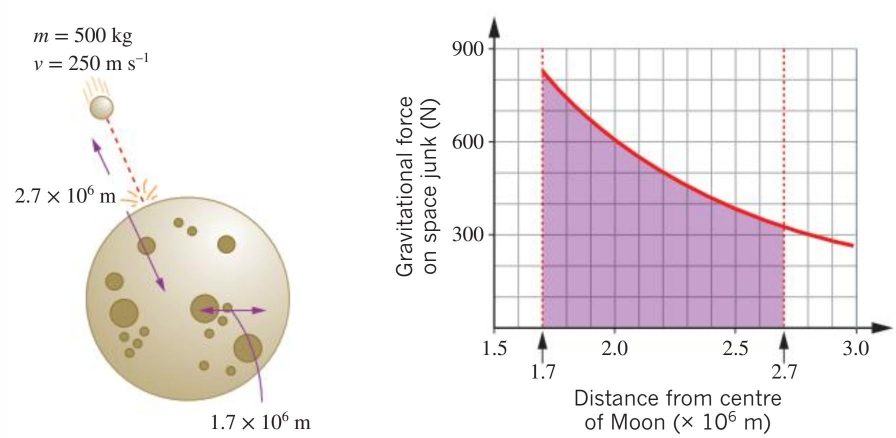


Thinking	Working
Count the number of shaded squares. (In this example, count the partially shaded squares as half squares.)	Number of shaded squares = 2
Calculate the area (energy value) of each square.	$E_{square} = 0.5 \times 10^7 \times 20$ $= 1 \times 10^8$ J
To calculate the energy change, multiply the number of shaded squares by the energy value of each square.	$\Delta E_g = 2 \times (1 \times 10^8)$ $= 2 \times 10^8$ J

Worked example: Try yourself 1.3.3

CHANGE IN GRAVITATIONAL POTENTIAL ENERGY USING A FORCE–DISTANCE GRAPH

A 500 kg lump of space junk is plummeting towards the Moon. The Moon has a radius of 1.7×10^6 m. Using the force–distance graph, determine the decrease in gravitational potential energy of the junk as it falls to the Moon's surface.



PHYSICSFILE

Using a force–distance graph in a constant field

A force–distance graph can also be used to calculate the change in gravitational potential energy of an object falling in a uniform gravitational field. Consider the graph in Figure 1.3.7 of a 10 kg rock that falls from a height of 40 m to 10 m.

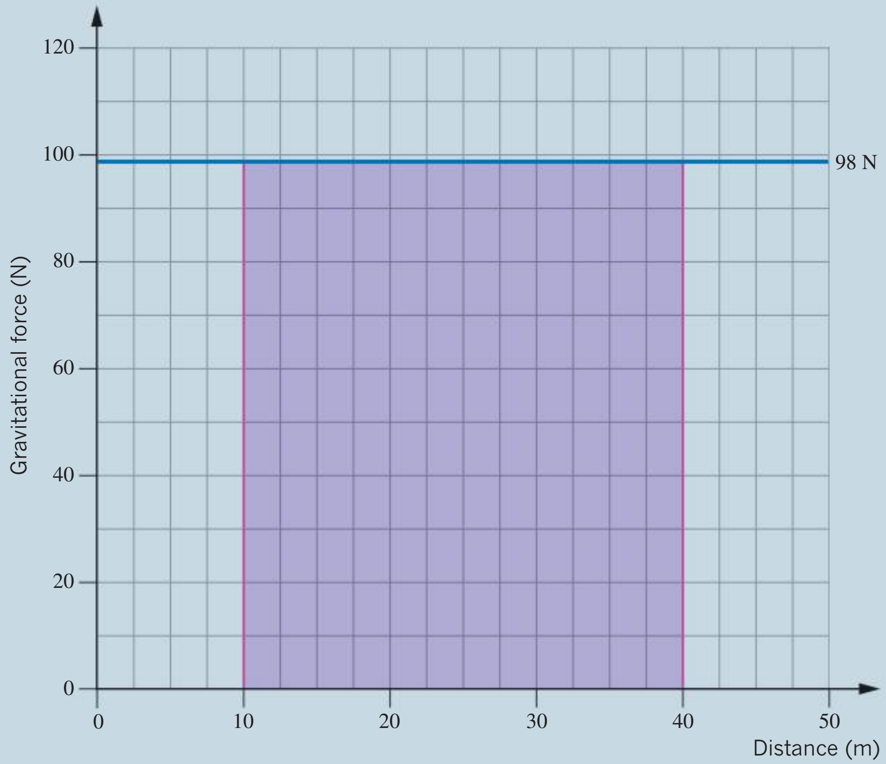


FIGURE 1.3.7 In a uniform field, the gravitational force–distance graph is a horizontal line.

The area under the graph in Figure 1.3.7 is 98 N × 30 m = 294 J, which is exactly what would have been calculated using the formula $E_g = mg\Delta h$. Therefore, it is more convenient to use the formula in uniform field situations.

USING THE GRAVITATIONAL FIELD STRENGTH–DISTANCE GRAPH

The change in gravitational potential energy of an object can also be calculated using a graph of the gravitational field strength of an object, as shown in Figure 1.3.8.

The area under a gravitational field strength–distance graph gives a quantity that has units of N kg^{-1} × m, which is equivalent to J kg^{-1}, so the area indicates the change in energy for each kilogram of the object's mass. To find the work done or energy change (J), the area (J kg^{-1}) must therefore be multiplied by the mass (kg) of the object.

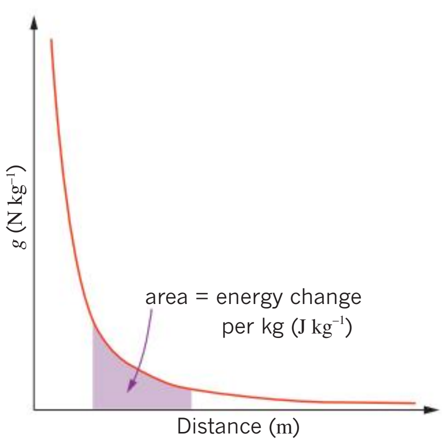


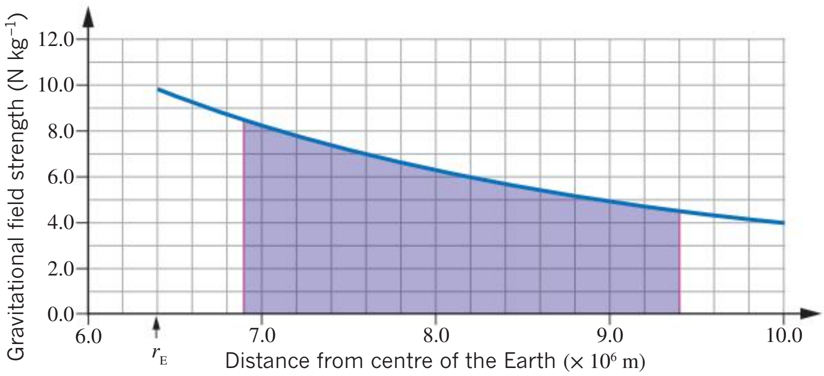
FIGURE 1.3.8 A graph of gravitational field strength–distance can also be used to analyse the energy changes of a body moving through a gravitational field.

The area under a gravitational field–distance graph gives the energy change per kilogram of mass. To find the change in energy, the area must be multiplied by the mass of the object in kg.

Worked example 1.3.4

CHANGE IN GRAVITATIONAL POTENTIAL ENERGY USING A GRAVITATIONAL FIELD STRENGTH–DISTANCE GRAPH

A wayward satellite of mass 1500 kg has developed a highly elliptical orbit around the Earth. At its closest approach (perigee), it is just 500 km above the Earth's surface. Its furthest point (apogee) is 3000 km from the Earth's surface. Using the graph of the gravitational field strength of the Earth shown below, determine the approximate change in the gravitational potential energy of the satellite as it orbits. (Note: the radius of the Earth is 6400 km.)

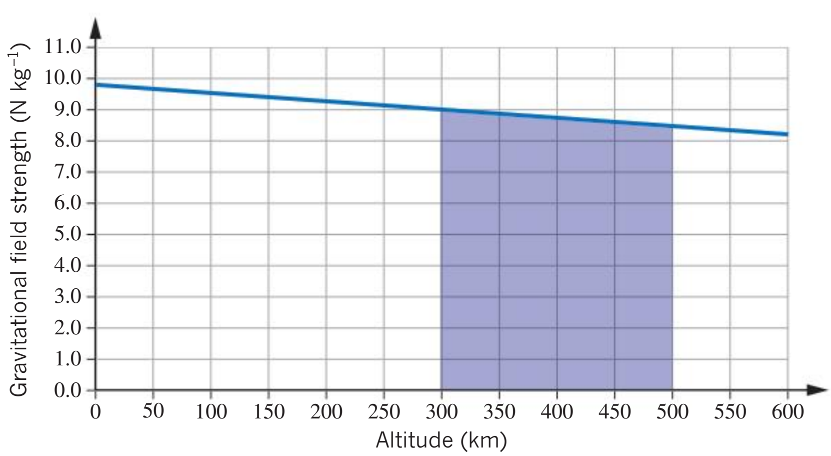


Thinking	Working
Count the number of shaded squares. Only count squares that are at least 50% shaded.	Number of shaded squares = 82
Calculate the energy value of each square.	$E_{square} = 0.2 \times 10^6 \text{ m} \times 1 \text{ N kg}^{-1}$ $= 2 \times 10^5 \text{ J kg}^{-1}$
To calculate the energy change, multiply the number of shaded squares by the energy value of each square and the mass of the satellite.	$\Delta E_g = 82 \times (2 \times 10^5) \times 1500$ $= 2.5 \times 10^{10} \text{ J}$

Worked example: Try yourself 1.3.4

CHANGE IN GRAVITATIONAL POTENTIAL ENERGY USING A GRAVITATIONAL FIELD STRENGTH–DISTANCE GRAPH

A 3000 kg Soyuz rocket moves from an orbital height of 300 km above the Earth's surface to dock with the International Space Station at a height of 500 km. Use the graph of the gravitational field strength of the Earth below to determine the approximate change in the gravitational potential energy of the rocket.



1.3 Review

SUMMARY

- The principles of work and conservation of energy are useful for understanding gravitational potential energy. This includes the following formulas:
  $W = Fs$
  $W = \Delta E$
  $E_k = \frac{1}{2}mv^2$
- The gravitational potential energy formula $E_g = mg\Delta h$ assumes that the Earth's gravitational field is constant. This is approximately true for objects that are within a few kilometres of the Earth's surface.
- The strength of the Earth's gravitational field decreases as altitude increases.
- The area under a gravitational force–distance graph gives the change in kinetic energy or change in gravitational potential energy of a body, and indicates the work done by the gravitational field.
- The area under a gravitational field–distance graph gives the change in energy per kilogram ($J\ kg^{-1}$) of the object. To calculate the energy change, the area is multiplied by the mass (kg).

KEY QUESTIONS

1 Which one of the following statements is correct? A satellite in a stable circular orbit around the Earth will have:

A varying potential energy as it orbits

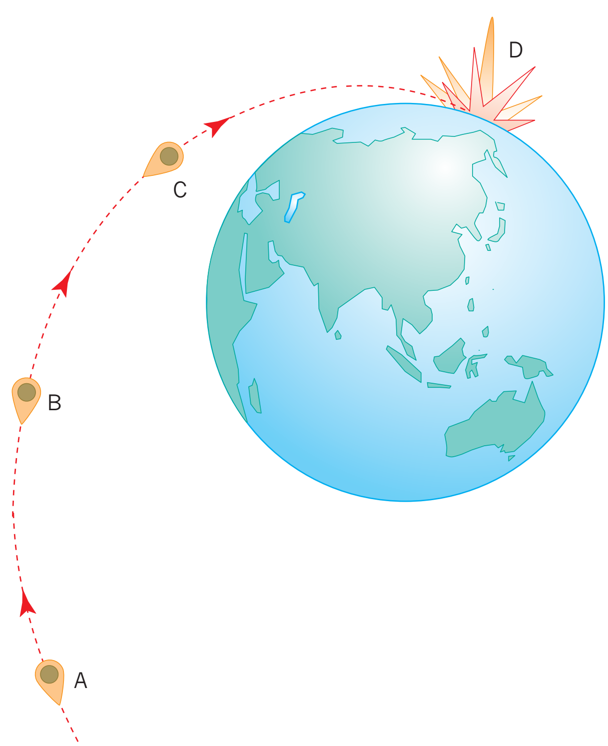
B varying kinetic energy as it orbits

C constant kinetic energy and constant potential energy

The following information applies to questions 2–4.

Ignore air resistance when answering these questions.

The path of a meteor plunging towards the Earth is as shown.



2 How does the gravitational field strength of the Earth change from point A to point D?

3 How will the acceleration of the meteor change as it travels along the path shown?

4 Which one or more of the following statements is correct?

A The kinetic energy of the meteor increases as it travels from A to D.

B The gravitational potential energy of the meteor decreases as it travels from A to D.

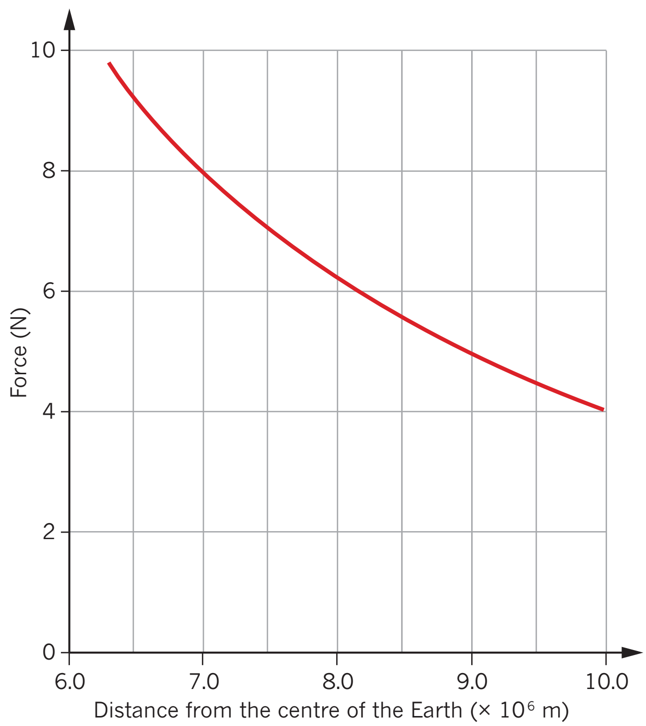
C The total energy of the meteor remains constant.

D The total energy of the meteor increases.

5 The *Saturn V* rocket that took the first astronauts to the Moon had a mass of 3000 tonnes. Its Stage I rockets fired for 6 minutes and took the rocket to an altitude of 67 km. How much work did the Stage I rockets do in this time?

6 The Valles Marineris on Mars is one of the most spectacular land formations in the solar system: a gigantic canyon 4000 km long, 200 km wide and 7 km deep. If a Martian explorer were to drop a 400 g rock from the edge of the canyon to its floor 7000 m below, how fast would the rock be going when it hit the bottom? The gravitational field strength on Mars is weaker than on Earth: $6.1\ N\ kg^{-1}$.

The following information applies to questions 7–9.
The graph shows the force on a mass of 1.0 kg as a function of its distance from the centre of the Earth.



7 **a** Use the graph to determine the gravitational force between the Earth and a 1.0 kg mass 100 km above the Earth's surface.

b Use the graph to determine the height above the Earth's surface at which a 1.0 kg mass would experience a gravitational force of 5.0 N.

8 A meteor of mass 1.0 kg is speeding towards the Earth. When this meteor is at a distance of 9.5×10^6 m from the centre of the planet, its speed is 4.0 km s^{-1}.

a Determine the kinetic energy of the meteor when it is 9.5×10^6 m from the centre of the Earth.

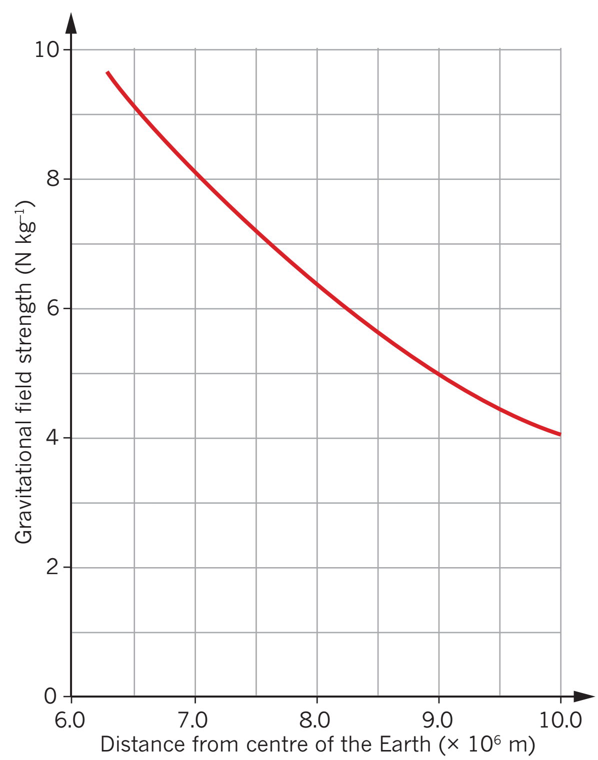
b Calculate the increase in kinetic energy of the meteor as it moves from a distance of 9.5×10^6 m from the centre of the Earth to a point that is 6.5×10^6 m from the centre.

c Ignoring air resistance, what is the kinetic energy of the meteor when it is 6.5×10^6 m from the centre of the Earth?

d How fast is the meteor travelling when it is 6.5×10^6 m from the centre of the Earth?

9 A communications satellite of mass 240 kg is launched from a space shuttle that is in orbit 600 km above the Earth's surface. The satellite travels directly away from the Earth and reaches a maximum distance of 8000 km from the centre of the Earth before stopping due to the influence of the Earth's gravitational field.
Use the graph to estimate the kinetic energy of this satellite as it was launched.

10 A 20 tonne remote-sensing satellite is in a circular orbit around the Earth at an altitude of 600 km. The satellite is moved to a new stable orbit with an altitude of 2600 km. Use the following graph to estimate the increase in the gravitational potential energy of the satellite as it moved from its lower orbit to its higher orbit.



Chapter review

01

KEY TERMS

acceleration due to gravity
altitude
apparent weight
field
gravimeter
gravitational constant
gravitational field
gravitational field strength
gravitational force
gravitational potential energy
inverse square law
Newton's law of universal gravitation
normal reaction force
torsion balance
uniform
weight

1 Use Newton's law of universal gravitation to calculate the gravitational force acting on a person with a mass of 75 kg. Use the following data:
$m_{Earth} = 6.0 \times 10^{24}$ kg
$r_{Earth} = 6400$ km

2 The gravitational force of attraction between Saturn and Dione, a moon of Saturn, is equal to 2.79×10^{20} N. Calculate the orbital radius of Dione. Use the following data:
mass of Dione = 1.05×10^{21} kg
mass of Saturn = 5.69×10^{26} kg

3 Of all the planets in the solar system, Jupiter exerts the largest force on the Sun: 4.2×10^{23} N. Calculate the acceleration of the Sun due to this force, using the following data: $m_{Sun} = 2.0 \times 10^{30}$ kg.

4 The planet Jupiter and the Sun exert gravitational forces on each other.
a Compare, qualitatively, the force exerted on Jupiter by the Sun to the force exerted on the Sun by Jupiter.
b Compare, qualitatively, the acceleration of Jupiter caused by the Sun to the acceleration of the Sun caused by Jupiter.

5 Calculate the acceleration due to gravity on the surface of Mars if it has a mass of 6.4×10^{23} kg and a radius of 3400 km.

6 Calculate the apparent weight of a 50 kg person in an elevator under the following circumstances.
a accelerating downwards at 0.6 m s^{-2}
b moving downwards at a constant speed of 2 m s^{-1}

7 A comet of mass 1000 kg is plummeting towards Jupiter. Jupiter has a mass of 1.90×10^{27} kg and a planetary radius of 7.15×10^{7} m. If the comet is about to crash into Jupiter, calculate the:
a magnitude of the gravitational force that Jupiter exerts on the comet
b magnitude of the gravitational force that the comet exerts on Jupiter
c acceleration of the comet towards Jupiter
d acceleration of Jupiter towards the comet.

8 A person standing on the surface of the Earth experiences a gravitational force of 900 N. What gravitational force will this person experience at a height of two Earth radii above the Earth's surface?
A 900 N
B 450 N
C zero
D 100 N

9 During a space mission, an astronaut of mass 80 kg initially accelerates at 30 m s^{-2} upwards, then travels in a stable circular orbit at an altitude where the gravitational field strength is 8.2 N kg^{-1}.
a What is the apparent weight of the astronaut during lift-off?
A zero
B 780 N
C 2400 N
D 3200 N
b During the lift-off phase, the astronaut will feel:
A lighter than usual
B heavier than usual
C the same as usual
c The true weight of the astronaut during the lift-off phase is:
A lower than usual
B greater than usual
C the same as usual
d During the orbit phase, the apparent weight of the astronaut is:
A zero
B 780 N
C 2400 N
D 660 N
e During the orbit phase, the true weight of the astronaut is:
A zero
B 780 N
C 2400 N
D 660 N

10 What are the main steps to follow when drawing gravitational field lines?

11 A set of bathroom scales is calibrated so that when the person standing on it has a weight of 600 N, the scales read 61.5 kg. What gravitational field strength has been assumed in this setting?

12 The Earth is a flattened sphere. Its radius at the poles is 6357 km compared to 6378 km at the equator. The Earth's mass is 5.97×10^{24} kg.

a Calculate the Earth's gravitational field strength at the equator.

b Using the information in part (a), calculate how much stronger the gravitational field would be at the North Pole compared with the equator. Give your answer as a percentage of the strength at the equator.

13 Neptune has a planetary radius of 2.48×10^7 m and a mass of 1.02×10^{26} kg.

a Calculate the gravitational field strength on the surface of Neptune.

b A 250 kg lump of ice is falling directly towards Neptune. What is its acceleration as it nears the surface of Neptune? Ignore any drag effects.

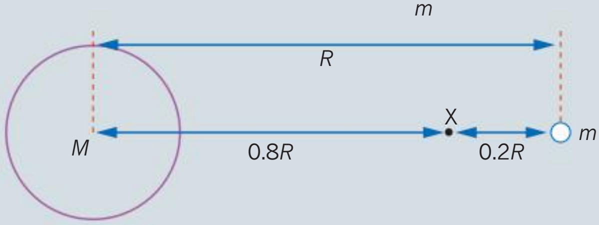
A 9.8 m s^{-2}

B zero

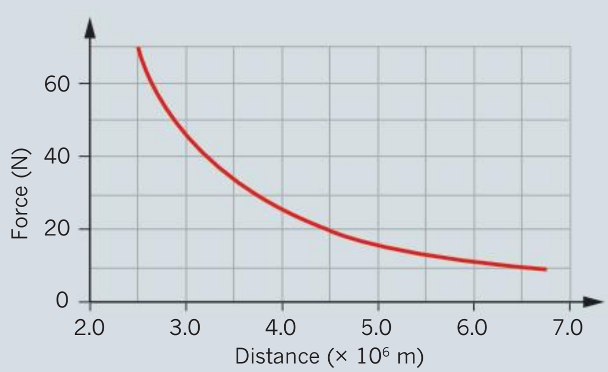
C 11 m s^{-2}

D 1.6 m s^{-2}

14 Two stars of masses *M* and *m* are in orbit around each other. As shown in the following diagram, they are a distance *R* apart. A spacecraft located at point X experiences zero net gravitational force from these stars. Calculate the value of the ratio $\frac{M}{m}$.



15 A 20 kg rock is speeding towards Mercury. The following graph shows the force on the rock as a function of its distance from the centre of the planet. The radius of Mercury is 2.4×10^6 m.



When the rock is 3.0×10^6 m from the centre of the planet, its speed is estimated at 1.0 km s^{-1}. Using the graph, estimate the:

a increase in kinetic energy of the rock as it moves to a point that is just 2.5×10^6 m from the centre of Mercury

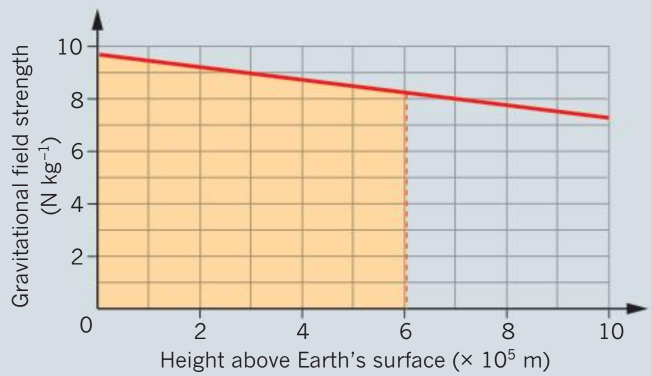
b kinetic energy of the rock at this closer point

c speed of the rock at this point

d gravitational field strength at 2.5×10^6 m from the centre of Mercury.

The following information relates to questions 16–20.

The diagram shows the gravitational field and distance near the Earth. A wayward satellite of mass 1000 kg is drifting towards the Earth.



16 What is the gravitational field strength at an altitude of 300 km?

17 Which of the following units is associated with the area under this graph?

A J

B m s^{-2}

C J s

D J kg^{-1}

18 Which one of the following quantities is represented by the shaded area on the graph? (Ignore air resistance.)

A the kinetic energy per kilogram of the satellite at an altitude of 600 km

B the loss in gravitational potential energy of the satellite

C the loss in gravitational potential energy per kilogram of the satellite as it falls to the Earth's surface

D the increase in gravitational potential energy of the satellite as it falls to the Earth's surface

19 How much kinetic energy does the satellite gain as it travels from an altitude of 600 km to an altitude of 200 km?

20 In reality, would the satellite gain the amount of kinetic energy that you have calculated in Question 19? Explain your answer.

CHAPTER

02 Electric and magnetic fields

In 1820, Hans Christian Oersted discovered that an electric current could produce a magnetic field. His work established the initial ideas behind electromagnetism. Since then, our understanding and application of electromagnetism has developed to the extent that much of our modern way of living relies upon it.

In this chapter you will investigate electric and magnetic fields, the concepts that apply to each, and some of the interactions between these closely related phenomena.

Key knowledge

By the end of this chapter you will have studied the physics of electric and magnetic fields, and will be able to:

- describe magnetism and electricity using a field model
- investigate and compare theoretically and practically gravitational, magnetic and electric fields, including directions and shapes of fields, attractive and repulsive fields, and the existence of dipoles and monopoles
- investigate and compare electrical fields about a point charge (positive or negative) with reference to:
  - the direction of the field
  - the shape of the field
  - the use of the inverse square law to determine the magnitude of the field
  - potential energy changes (qualitative) associated with a point charge moving in the field
- investigate and apply theoretically and practically a vector field model to magnetic phenomena, including shapes and directions of fields produced by bar magnets and by current-carrying wires, loops and solenoids
- identify fields as static or changing, and as uniform or non-uniform
- analyse the use of an electric field to accelerate a charge, including:
  - electric field and electric force concepts: $E = k\frac{Q}{r^2}$ and $F = k\frac{q_1q_2}{r^2}$
  - potential energy changes in a uniform electric field: $W = qV$ and $E = \frac{V}{d}$
  - the magnitude of the force on a charged particle due to a uniform electric field: $F = qE$
- analyse the use of a magnetic field to change the path of a charged particle, including:
  - the magnitude and direction of the force applied to an electron beam by a magnetic field: $F = qvB$, in cases where the directions of v and B are perpendicular or parallel
- describe the interaction of two fields, allowing that electric charges, magnetic poles and current carrying conductors can either attract or repel, whereas masses only attract each other
- investigate and analyse theoretically and practically the force on a current-carrying conductor due to an external magnetic field, $F = nIlB$, where the directions of I and B are either perpendicular or parallel to each other.

2.1 Electric fields

FIGURE 2.1.1 Charged plasma follows lines of the electric field produced by a Van de Graaff generator.

FIGURE 2.1.2 The gravitational field of the Earth applies a force on the skydiver, while the gravitational field of the skydiver exerts a force on the Earth.

FIGURE 2.1.3 The girl's hair follows the lines of the electric field produced when she became charged while sliding down a plastic slide.

A field is a region of space where objects experience a force due to a physical property related to the field. Gravity, electricity and magnetism can all be described by fields. In Chapter 1 'Gravity', the direction, shape and strength of gravitational fields around a mass were described. In this section the electric field will be explained.

An **electric field** surrounds positive and negative charges, and exerts a force on other charges within the field. Just as a gravitational field can be represented by field lines, so can the electric field around charged objects. This is shown in Figure 2.1.1.

ELECTRIC FIELDS

There are four fundamental forces in nature that act at a distance. That is, they can exert a force on an object without making any physical contact with it. These are called non-contact forces, and include the strong nuclear force, the weak nuclear force, the electromagnetic force and the gravitational force.

In order to understand these forces, scientists use the idea of a field. A field is a region of space around an object that has certain physical properties, such as electric charge or mass. Another object with that physical property in the field will experience a force without any contact between the two objects.

For example, as you saw in Chapter 1, there is a gravitational field around the Earth due to the mass of the Earth. Any object with mass that is located within this gravitational field experiences a force of attraction towards the Earth. According to Newton's third law, there is also an equal and opposite force due to gravity acting on the Earth due to the gravitational field of the object. An example of this is shown in Figure 2.1.2.

Similarly, any charged object has a region of space around it (an electric field) where another charged object will experience a force. This is one aspect of the electromagnetic force. Unlike gravity, which only exerts an attractive force, electric fields can exert forces of attraction or repulsion.

ELECTRIC FIELD LINES

An electric field is a vector quantity, which means it has both direction and strength.

In order to visualise electric fields around charged objects you can use electric **field lines**. Some field lines are already visible—for example the girl's hair in Figure 2.1.3 is tracing out the path of the field lines. Diagrams of field lines can also be constructed.

Field lines are drawn with arrowheads on them indicating the direction of the force that a small positive test charge would experience if it were placed in the electric field. Therefore, field lines point away from positively charged objects and towards negatively charged objects. Usually, only a few representative lines are drawn.

> Remember: like charges repel and unlike (opposite) charges attract.

The density of field lines (how close they are together) is an indication of the relative strength of the electric field. This is explained in more detail later in this section.

Rules for drawing electric field lines

When drawing electric field lines (in two dimensions) around a charged object there are a few rules that need to be followed.

- Electric field lines go from positively charged objects to negatively charged objects.
- Electric field lines start and end at 90° to the surface, with no gap between the lines and the surface.
- Field lines can never cross; if they did it would indicate that the field is in two directions at that point, which can never happen.

- Around small charged spheres, called **point charges**, the field lines radiate like spokes on a wheel.
- Around point charges you should draw at least eight field lines: top, bottom, left, right and another field line in between each of these.
- Between two point charges, the direction of the field at any point is the resultant field vector determined by adding the field vectors due to each of the two point charges.
- Between two oppositely charged parallel plates, the field lines between the plates are evenly spaced and are drawn straight from the positive plate to the negative plate.
- Always remember that these drawings are two-dimensional representations of a three-dimensional field.

  Figure 2.1.4 shows some examples of how to draw electric field lines.

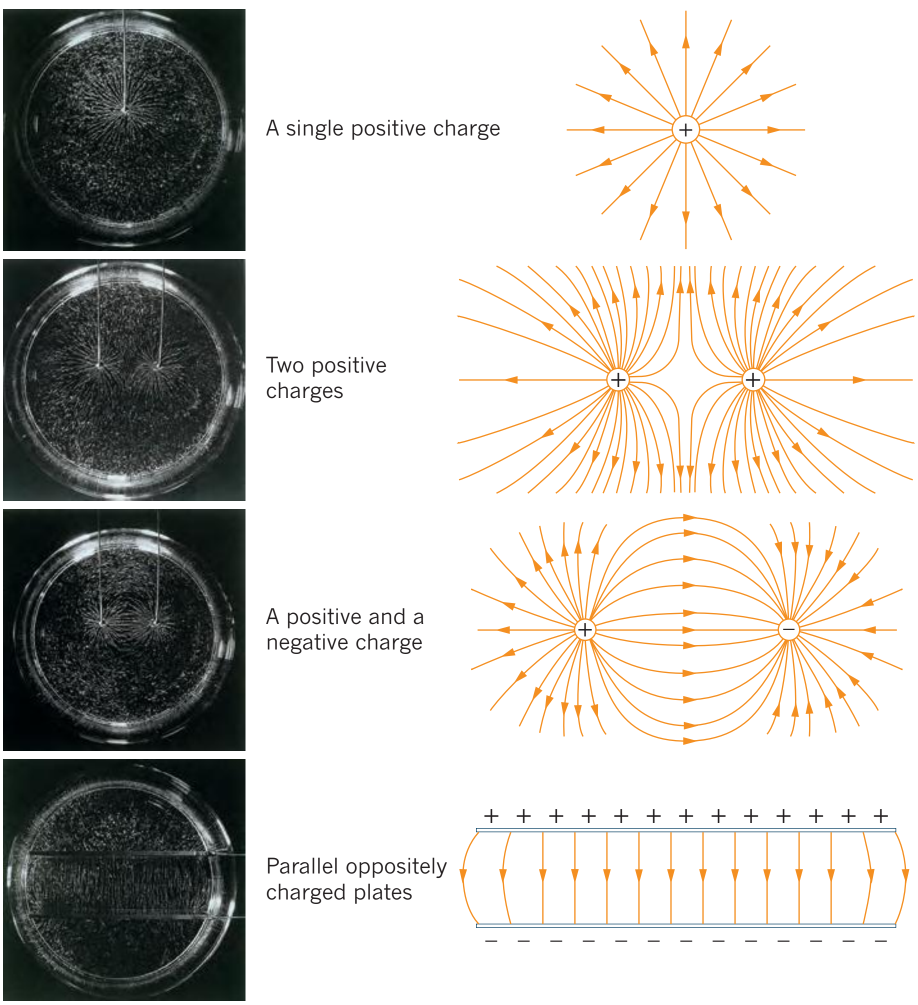


FIGURE 2.1.4 Grass seeds suspended in oil align themselves with the electric field. The diagram next to each photo shows lines representing the electric field.

Strength of the electric field

The distance between adjacent field lines indicates the strength of the field. Around a point charge, the field lines are closer together near the charge and get further apart as you move further away. You can see this in the field-line diagrams in Figure 2.1.4. Therefore, the **electric field strength** decreases as the distance from a point charge increases.

A uniform electric field is established between two parallel metal plates that are oppositely charged. The field strength is constant at all points within a uniform electric field, so the field lines are parallel.

FORCES ON FREE CHARGES IN ELECTRIC FIELDS

If a charged particle, such as an electron, were placed within an electric field, it would experience a force. The direction of the field and the sign of the charge allow you to determine the direction of the force.

Figure 2.1.5 shows a positive test charge (proton) and a negative test charge (electron), within a uniform electric field. Recall that the direction of an electric field is defined as the direction of the force that a positive charge would experience within the electric field. So, an electron will experience a force in the opposite direction to the electric field, while a proton will experience a force in the same direction as the field.

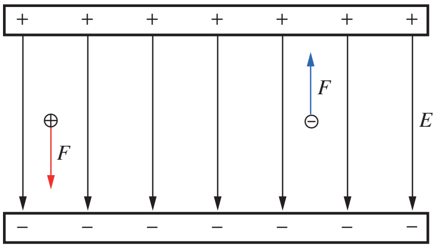


FIGURE 2.1.5 The direction of the electric field (E) indicates the direction in which a force would act on a positive charge. A negative charge would experience a force in the opposite direction to the field.

PHYSICSFILE

Bees

Bees are thought to use electric fields to communicate, find food and to avoid flowers that have been visited by another bee recently. Their antennae are bent (deflected) by electric fields and they can sense the amount of deflection (see Figure 2.1.6).

The charge that builds up on their bodies helps them collect pollen grains and transport them to other flowers. The altered electric field around a flower that has recently been visited is a signal to other bees to find food elsewhere.

FIGURE 2.1.6 A bee can detect changes in the electric field around its body.

The magnitude of the force experienced by a charged particle due to an electric field can be determined using the equation:

$F = qE$

where F is the force on the charged particle (N)

q is the charge of the object experiencing the force (C)

E is the strength of the electric field (N C^{-1})

This equation illustrates that the force experienced by a charge is proportional to the strength of the electric field, E, and the size of the charge, q. The force on the charged particle will cause the charged particle to accelerate in the field. This means that the particle could increase its velocity, decrease its velocity, or change its direction while in the field.

To calculate the acceleration due to the force experienced, you can use the equation from Newton's second law:

$F = ma$

where m is the mass of the accelerating particle (kg)

a is the acceleration (m s^{-2}).

Worked example 2.1.1

USING $F = qE$

Calculate the magnitude of the uniform electric field that would cause a force of 5.00×10^{-21} N on an electron. ($q_e = -1.602 \times 10^{-19}$ C)	
Thinking	**Working**
Rearrange the relevant equation to make electric field strength the subject.	$F = qE$ $E = \frac{F}{q}$
Substitute the values for F and q into the rearranged equation and calculate the answer. (As only magnitude is required, q can be kept positive.)	$E = \frac{F}{q}$ $= \frac{5.00 \times 10^{-21}}{1.602 \times 10^{-19}}$ $= 3.12 \times 10^{-2}$ N C^{-1}

Worked example: Try yourself 2.1.1

USING $F = qE$

Calculate the magnitude of the uniform electric field that creates a force of 9.00×10^{-23} N on a proton.

($q_p = +1.602 \times 10^{-19}$ C)

Electric field strength

Electric field strength can be thought of as the force applied per coulomb of charge, which is expressed in the equation:

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

An alternative measure of the electric field strength is volts per metre, which is calculated using the equation:

$$E = \frac{V}{d}$$

where d is distance (m).

You can equate both expressions and rearrange them to find an expression for the work done (J) to make a charged particle move a distance against a potential difference:

$$\frac{F}{q} = \frac{V}{d}$$

$Fd = qV$ and since $W = Fd$

$$W = qV$$

EXTENSION

Gravitational force and electric force

Oppositely charged parallel plates can be arranged one above the other, such that the electric field is vertical. The direction of the field can then be manipulated to create an upwards force on a charged particle in the field.

If the electric force created by the field on the charged object is equal to the gravitational force on (or weight of) the object, then these two forces can add to provide a net force equal to zero. This means that the charged object will either be suspended between the plates, or (by Newton's first law of motion) will be falling or rising at constant velocity.

This phenomenon was used by Robert Millikan and his PhD student Harvey Fletcher in their oil drop experiment, performed in 1909, to determine the fundamental charge of an electron to within 1% of the currently accepted value.

WORK DONE IN UNIFORM ELECTRIC FIELDS

Electrical potential energy is a form of energy that is stored in a field. Work is done on the field when a charged particle is forced to move in the electric field. Conversely, when energy is stored in the electric field then work can be done by the field on the charged particle.

Electrical potential (V) is defined as the work required per unit charge to move a positive point charge from infinity to a place in the electric field. The electrical potential at infinity is defined as zero. This definition leads to the equation:

$$V = \frac{W}{q}$$

$$W = qV$$

where W is the work done on a positive point charge or on the field (J)

q is the charge of the point charge (C)

V is the electrical potential (J C^{-1}) or volts (V)

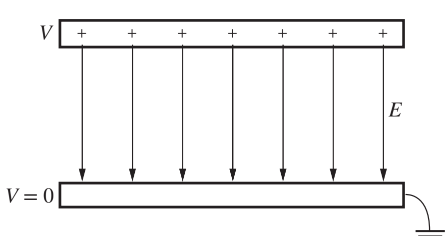


FIGURE 2.1.7 The potential of two plates when one has a positive potential and the other is earthed.

Consider two parallel plates, as shown in Figure 2.1.7, in which the positive plate is at a potential (V) and the other plate is earthed, which is defined as zero potential. The difference in potential between these two plates is called the electrical potential difference (V).

Between any two points in an electric field (E) separated by a distance (d) that is parallel to the field, the potential difference V is then defined as the change in the electrical potential between these two points. See Figure 2.1.8.

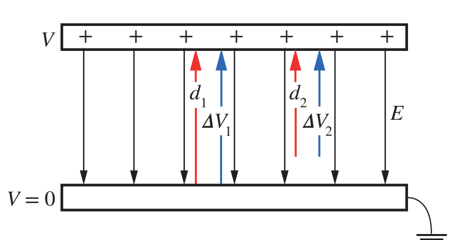


FIGURE 2.1.8 The potential difference between two points in a uniform electric field.

$$E = \frac{V}{d}$$

$$V = Ed$$

where V is the difference in electrical potential (V)

E is the electrical field strength (V m^{-1})

d is the distance between points, parallel to the field (m)

CALCULATING WORK DONE

By combining the two equations mentioned so far, you can derive an equation for calculating the work done on a point test charge to move it a distance across a potential difference.

$W = qV$ and $V = Ed$

so $W = qEd$

where W is the work done on the point charge or on the field (J)

q is the charge of the point charge (C)

E is the electrical field strength (V m^{-1} or N C^{-1})

d is the distance between points, parallel to the field (m).

EXTENSION

Dimensional analysis of the units for field strength

For the value of E in the above equation you can use either units for electrical field strength, since they are equivalent. The following dimensional analysis shows why they are equivalent.

$E = \frac{V}{d}$ and $E = \frac{F}{q}$

so $\frac{V}{d} = \frac{F}{q}$

Looking at the units for each side of the equation, V m^{-1} must equal N C^{-1}.

To prove this, you can break down each unit:

V = J C^{-1} = kg m^2 s^{-2} C^{-1}

so V m^{-1} = (kg m^2 s^{-2} C^{-1}) m^{-1} = kg m s^{-2} C^{-1}

Since N = kg m s^{-2}:

V m^{-1} = kg m s^{-2} C^{-1} = N C^{-1}

Work done by or on an electric field

When calculating work done, which changes the electrical potential energy, remember that work can be done either:

- by the electric field on a charged object, or
- on the electric field by forcing the object to move.

You need to examine what's happening in a particular situation to know how the work is being done.

For example, if a charged object is moving in the direction it would naturally tend to go within an electric field, then work is done by the field. So when a positive point charge is moved in the direction of the electric field, the electric field has done work on the point charge. (Refer to q_2 in Figure 2.1.9.)

When work is done by a charged object on an electric field, the object is forced to move against the direction it would naturally go. Work has been done on the field by forcing the object to move. For example, if a force causes a positive charge to move towards the positive plate within a uniform electric field, work has been done on the electric field by forcing the object to move. (See q_1 in Figure 2.1.9.)

If a charge doesn't move any distance parallel to the field then no work is done on or by the field. (See q_3 in Figure 2.1.9.)

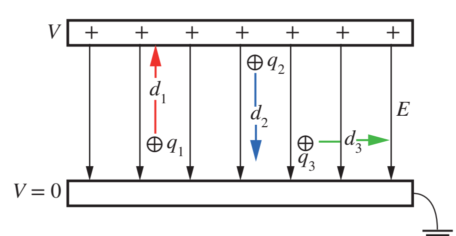


FIGURE 2.1.9 Work is being done on the field by moving q_1 and work is being done by the field on q_2. No work is done on q_3 since it is moving perpendicular to the field.

Worked example 2.1.2

WORK DONE ON A CHARGE IN A UNIFORM ELECTRIC FIELD

A student sets up a parallel plate arrangement so that one plate is at a potential of 12.0 V and the other earthed plate is positioned 0.50 m away. Calculate the work done to move a proton a distance of 10.0 cm towards the negative plate.

($q_p = +1.602 \times 10^{-19}$ C)

In your answer identify what does the work and what the work is done on.

Thinking	Working
Identify the variables presented in the problem to calculate the electric field strength *E*.	$V_2 = 12.0$ V $V_1 = 0$ V $d = 0.50$ m
Use the equation $E = \frac{V}{d}$ to determine the electric field strength.	$E = \frac{V}{d}$ $= \frac{12.0 - 0}{0.50}$ $= 24.0$ V m^{-1}
Use the equation $W = qEd$ to determine the work done. Note that *d* here is the distance that the proton moves.	$W = qEd$ $= 1.602 \times 10^{-19} \times 24.0 \times 0.100$ $= 3.84 \times 10^{-19}$ J
Determine if work is done on the charge by the field or if work is done on the field.	As the positively charged proton is moving naturally towards the negative plate then work is done on the proton by the field.

Worked example: Try yourself 2.1.2

WORK DONE ON A CHARGE IN A UNIFORM ELECTRIC FIELD

A student sets up a parallel plate arrangement so that one plate is at a potential of 36.0 V and the other earthed plate is positioned 2.00 m away. Calculate the work done to move an electron a distance of 75.0 cm towards the negative plate.

($q_e = -1.602 \times 10^{-19}$ C)

In your answer identify what does the work and what the work is done on.

2.1 Review

SUMMARY

- An electric field is a region of space around a charged object in which another charged object will experience a force.
- Electric fields are represented using field lines.
- Electric field lines point in the direction of the force that a positive charge within the field would experience.
- A positive charge experiences a force in the direction of the electric field and a negative charge experiences a force in the opposite direction to the field.
- The spacing between the field lines indicates the strength of the field. The closer together the lines are, the stronger the field.
- Electric field strength can be expressed as $E = \frac{F}{q}$ and also $E = \frac{V}{d}$
- Around point charges the electric field radiates in all directions (three dimensionally).
- Between two oppositely charged parallel plates, the field lines are parallel and therefore the field has a uniform strength.
- When charges are in an electric field, they accelerate in the direction of the force acting on them.
- The force on a charged particle can be determined using the equation $F = qE$.

- Force can be related to the acceleration of a particle using the equation $F = ma$.
- Electrical potential energy is stored in an electric field.
- When a charged object is moved against the direction it would naturally move in an electric field, then work is done on the field.
- When a charged object is moved in the direction it would naturally tend to move in an electric field, then the field does work on the particle.
- The work done on or by an electric field can be calculated using the equations $W = qV$ or $W = qEd$.

KEY QUESTIONS

1 Which of the following options correctly describes an electric field?
   A a region around a charged object that causes a charge on other objects within that region
   B a region around a charged object that causes a force on other objects within that region
   C a region around a charged object that causes a force on other charged objects in that region
   D a region around an object that causes a force on other objects within that region

2 Which of the following options correctly defines the direction of an electric field?
   A away from a negatively charged object
   B away from a positively charged object
   C away from a neutrally charged object
   D towards a positively charged object

3 Identify whether the rules below for drawing electric field lines are true or false:
   a Electric field lines start and end at 90° to the surface, with no gap between the lines and the surface.
   b Field lines can cross; this indicates that the field is in two directions at that point.
   c Electric fields go from negatively charged objects to positively charged objects.
   d Around small charged spheres called point charges you should draw at least eight field lines: top, bottom, left, right and in between each of these.
   e Around point charges the field lines radiate like spokes on a wheel.
   f Between two point charges the direction of the field at any point is the field due to the closest of the two point charges.
   g Between two oppositely charged parallel plates the field between the plates is evenly spaced and is drawn straight from the negative plate to the positive plate.

4 Calculate the force applied to a balloon carrying a charge of 5.00 mC in a uniform electric field of 2.50 N C^{-1}.

5 Calculate the charge on a plastic disk if it experiences a force of 0.025 N in a uniform electric field of 18 N C^{-1}.

6 Calculate the acceleration of an electron in a uniform electric field of 3.25 N C^{-1}. The mass of an electron is 9.11×10^{-31} kg and its charge is -1.602×10^{-19} C.

7 Calculate the potential difference that exists between two points separated by 30.0 cm, parallel to the field lines, in an electric field of strength 4000 V m^{-1}.

8 A researcher sees an oil drop with a of mass 1.161×10^{-14} kg stationary between two horizontal parallel plates. Between the plates exists an electric field of strength 3.55×10^4 N C^{-1}. The field is pointing vertically downwards. Calculate how many extra electrons are present on the oil drop. ($q_e = -1.602 \times 10^{-19}$ C and $g = 9.8$ N kg^{-1})

9 For each for each of the following charged objects in a uniform electric field, determine if work was done on the field or by the field or if no work is done.
   a An electron moves towards a positive plate.
   b A positively charged point remains stationary.
   c A proton moves towards a positive plate.
   d A lithium ion (Li^+) moves parallel to the plates.
   e An alpha particle moves away from a negative plate.
   f A positron moves away from a positive plate.

10 An alpha particle is located in a parallel plate arrangement that has a uniform electric field of 34.0 V m^{-1}.
   a Calculate the work done to move the alpha particle a distance of 1.00 cm from the earthed plate to the plate with a positive potential. ($q_\alpha = +3.204 \times 10^{-19}$ C)
   b For the situation in part (a) decide whether work was done on the field, by the field or if no work was done.

2.2 Coulomb's law

Electricity is one of nature's fundamental forces. It was Charles Coulomb, in 1785, who first published the quantitative details of the force that acts between two electric charges. The force between any combination of electrical charges can be understood in terms of the force between two 'point charges' separated by a certain distance, as seen in Figure 2.2.1. The effect of distance on the electric field strength from a single charge and the force created by that field between charges is explored in this section.

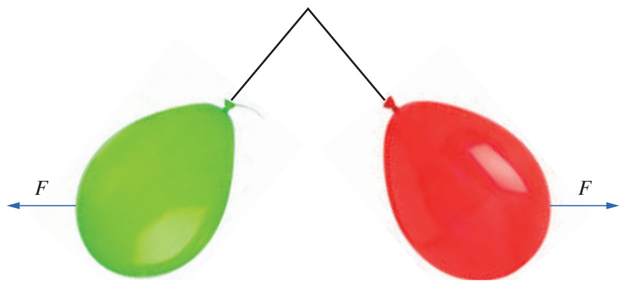


FIGURE 2.2.1 Two similarly charged balloons will repel each other by applying a force on each other.

THE FORCE BETWEEN CHARGED PARTICLES

Coulomb found that the force between two point charges (q_1) and (q_2) separated by a distance (r) was proportional to the product of the two charges, and inversely proportional to the square of the distance between them.

This is another example of an inverse square law, as discussed in Chapter 1.

Coulomb's law can be expressed by the following equation:

$$F = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r^2}$$

where F is the force on each charged object (N)

q_1 is the charge on one point (C)

q_2 is the charge on the other point (C)

r is distance between each charged point (m)

ε_0 is the permittivity of free space, which is equal to 8.8542×10^{-12} C^2 N^{-1} m^{-2}

By including the sign of the charges in the calculation, a positive force value indicates repulsion and a negative force value indicates attraction.

The permittivity of free space (ε_0) has a value of 8.8542×10^{-12} C^2 N^{-1} m^{-2} in air or a vacuum. As this value is constant for air or a vacuum, the expression at the front of Coulomb's law can be calculated. The result of the calculation is given the name Coulomb's constant (k) and is equal to 8.9875×10^9 N m^2 C^{-2}. For ease of calculations this is usually rounded to two significant figures as 9.0×10^9 N m^2 C^{-2}. So, in Coulomb's law, if:

$$k = \frac{1}{4\pi\varepsilon_0}$$

then the equation becomes:

$$F = k\frac{q_1 q_2}{r^2}$$

where $k = 9.0 \times 10^9$ N m^2 C^{-2}

Factors affecting the electric force

The force between two charged points is proportional to the product of the two charges as seen in Figure 2.2.2. If there was a force of 10 N between two charged points and either charge was doubled, then the force between the two points would increase to 20 N. It is interesting to note that regardless of the charge on each point, the forces on each point in a pair will be the same. For example, if q_1 is +10 μC and q_2 is +10 μC, then the repulsive force on each of these points would be equal in magnitude. The forces would also be equal on both points if q_1 is +100 μC and q_2 is +1 μC.

The force is also inversely proportional to the square of the distance between the two charged points. This means that if the distance between q_1 and q_2 is doubled, the force on each point charge will decrease to one-quarter of the previous value.

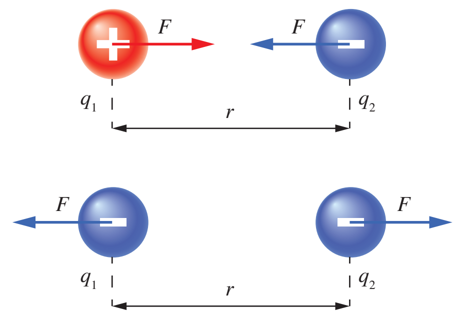


FIGURE 2.2.2 Forces acting between two point charges.

One coulomb in perspective

Using Coulomb's law you can calculate the force between two charges of 1 C each, placed 1 m apart. The force would be 9.0×10^9 N, or approximately 10^{10} N. This is equivalent to the weight provided by a mass of 918 000 tonnes (see Figure 2.2.3).

This demonstrates that a 1 C charge is a huge amount of charge. In reality, the amount of charge that can be placed on ordinary objects is a tiny fraction of a coulomb. Even a highly charged Van de Graaff generator will have only a few microcoulombs (1 μC = 10^{-6} C) of excess charge.

FIGURE 2.2.3 Two 1 C charges 1 m apart would produce a force of 10^{10} N, which is almost twice the weight of the Sydney Harbour Bridge.

Another way to get a feel for the magnitude of electrical forces is to realise that all matter is held together by the electrical forces between atoms. For example, the mass of Mount Everest is supported by the electrostatic repulsion between the electrons around neighbouring atoms in the rock underneath it. The strength of the hardest steel is due to the electrical forces of attraction between its ions and the delocalised electrons between them. In comparison to the Earth's gravitational force on an atom, the electrical forces between atoms are about a billion, billion (1×10^{18}) times stronger. In fact, only in the last stages of gravitational collapse of a giant star can the gravitational forces overwhelm the electrical forces between its atoms and cause the star to collapse into a super-dense neutron star.

Worked example 2.2.1

USING COULOMB'S LAW TO CALCULATE FORCE

Two small spheres A and B act as point charges separated by 10.0 cm in air. Calculate the force on each point charge if A has a charge of 3.00 μC and B has a charge of –45.0 nC.
(Use $\varepsilon_0 = 8.8542 \times 10^{-12}$ C^2 N^{-1} m^{-2}.)

Thinking	Working
Convert all values to SI units.	$q_A = 3.00 \times 10^{-6}$ C $q_B = 45.0 \times 10^{-9} = -4.50 \times 10^{-8}$ C $r = 0.100$ m
State Coulomb's law.	$F = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r^2}$
Substitute the values for q_A, q_B, r and ε_0 into the equation and calculate the answer.	$F = \frac{1}{4\pi \times 8.8542 \times 10^{-12}} \times \frac{3.00 \times 10^{-6} \times -4.50 \times 10^{-8}}{0.100^2}$ $= -0.121$ N
Assign a direction based on a negative force being attraction and a positive force being repulsion.	$F = 0.121$ N attraction

Worked example: Try yourself 2.2.1

USING COULOMB'S LAW TO CALCULATE FORCE

Two small spheres A and B act as point charges separated by 75.0 mm in air. Calculate the force on each point charge if A has a charge of 475 nC and B has a charge of 833 pC.
(Use $\varepsilon_0 = 8.8542 \times 10^{-12}$ C^2 N^{-1} m^{-2}.)

Worked example 2.2.2

USING COULOMB'S LAW TO CALCULATE CHARGE

Two small positive point charges with equal charge are separated by 1.25 cm in air. Calculate the charge on each point charge if there is a repulsive force of 6.48 mN between them.
(Use $k = 9.0 \times 10^9$ N m^2 C^{-2}.)

Thinking	Working
Convert all values to SI units.	$F = 6.48 \times 10^{-3}$ N $r = 1.25 \times 10^{-2}$ m
State Coulomb's law.	$F = k\frac{q_1 q_2}{r^2}$
Substitute the values for F, r and k into the equation and calculate the answer.	$q_1 q_2 = \frac{Fr^2}{k}$ $= \frac{6.48 \times 10^{-3} \times (1.25 \times 10^{-2})^2}{9.0 \times 10^9}$ $= 1.125 \times 10^{-16}$ Since $q_1 = q_2$: $q_1^2 = 1.125 \times 10^{-16}$ $q_1 = \sqrt{1.125 \times 10^{-16}}$ $= +1.06 \times 10^{-8}$ C $q_2 = +1.06 \times 10^{-8}$ C

Worked example: Try yourself 2.2.2

USING COULOMB'S LAW TO CALCULATE CHARGE

Two small point charges are charged by transferring a number of electrons from q_1 to q_2, and are separated by 12.7 mm in air. The charges the two points are equal and opposite. Calculate the charge on q_1 and q_2 if there is an attractive force of 22.5 μN between them. (Use $k = 9.0 \times 10^9$ N m^2 C^{-2}.)

THE ELECTRIC FIELD AT A DISTANCE FROM A CHARGE

In the previous section 'Electric fields', the electric field, E, is defined as being proportional to the force exerted on a positive test charge and inversely proportional to the magnitude of that charge, and is measured in N C^{-1}. Defining the electric field in this way means that it is independent of the size of the charge and describes only the effect of the charge creating the field at a particular point.

It is useful also to be able to determine the electric field at a distance from a single point charge:

The magnitude of the electric field at a distance from a single point charge is given by:

$$E = k\frac{Q}{r^2}$$

where E is the strength of the electric field around a point (N C^{-1})

Q is the charge on the point creating the field (C)

r is the distance from the charge (m)

$k = 9.0 \times 10^9$ N m^2 C^{-2}

The magnitude of E that is determined is independent of the value of the test charge and depends only on the charge, Q, producing the field. This formula can also be referred to as Coulomb's law, in this case for determining the magnitude of the electric field produced by a point charge.

Worked example 2.2.3

ELECTRIC FIELD OF A SINGLE POINT CHARGE

Calculate the magnitude and direction of the electric field at a point P at a distance of 20 cm below a negative point charge, Q, of 2.0×10^{-6} C.

Thinking	Working
Convert units to SI units as required.	$Q = -2.0 \times 10^{-6}$ C r = 20 cm = 0.20 m
Substitute the known values to find the magnitude of E using: $E = k\frac{Q}{r^2}$	$E = k\frac{Q}{r^2}$ $= 9.0 \times 10^9 \times \frac{2.0 \times 10^{-6}}{0.20^2}$ $= 4.5 \times 10^5$ N C^{-1}
The direction of the field is defined as that acting on a positive test charge (see previous section). Point P is below the charge.	Since the charge is negative, the direction will be toward the charge Q, or upwards.

Worked example: Try yourself 2.2.3

ELECTRIC FIELD OF A SINGLE POINT CHARGE

Calculate the magnitude and direction of the electric field at point P at a distance of 15 cm to the right of a positive point charge, Q, of 2.0×10^{-6} C.

2.2 Review

SUMMARY

- Coulomb's law for the force between two charges q_1 and q_2 separated by a distance of r is:
  $F = \frac{1}{4\pi\varepsilon_0}\frac{q_1q_2}{r^2}$
- The constant, ε_0, in Coulomb's law has a value of 8.8542×10^{-12} C^2 N^{-1} m^{-2}.
- For air or a vacuum, the expression $\frac{1}{4\pi\varepsilon_0}$ at the front of Coulomb's law can be simplified to the value of k, called Coulomb's constant, which has a value of approximately 9.0×10^9 N m^2 C^{-2}. So:
  $F = k\frac{q_1q_2}{r^2}$
- The magnitude of the electric field, E, at a distance r from a single point charge Q is given by:
  $E = k\frac{Q}{r^2}$
  where $k = 9.0 \times 10^9$ N m^2 C^{-2}

KEY QUESTIONS

1 Choose a response from each shaded box to correctly complete the following table summarising forces, charges and actions.

Force	q_1 charge	q_2 charge	Action
a positive	positive	positive / negative	attraction / repulsion
b negative	positive / negative	positive	attraction / repulsion
c positive	negative	positive / negative	attraction / repulsion
d negative	positive / negative	negative	attraction / repulsion

2 A point charge, Q, is moved from a position 30 cm away from a test charge to a position 15 cm from the same test charge. If the magnitude of the original electric field, E, was 6.0×10^3 N C^{-1}, what is the magnitude of the electric field at the new position?

A 3.0×10^3 N C^{-1}
B 6.0×10^3 N C^{-1}
C 12.0×10^3 N C^{-1}
D 24.0×10^3 N C^{-1}

3 A hydrogen atom consists of a proton and an electron separated by a distance of 53 pm (picometres). Calculate the magnitude and sign of the force applied to a proton carrying a charge of $+1.602 \times 10^{-19}$ C by an electron carrying a charge of -1.602×10^{-19} C. (1 pm = 1×10^{-12} m and $\varepsilon_0 = 8.8542 \times 10^{-12}$ C^2 N^{-1} m^{-2}.)

4 The electric field is being measured at point P at a distance of 5.0 cm from a positive point charge, Q, of 3.0×10^{-6} C. What is the magnitude of the field at P to two significant figures? (Use $k = 9 \times 10^9$ N m^2 C^{-2}.)

5 Calculate the magnitude of the force that would exist between two point charges of 1.00 C each, separated by 1.00 km. (Use $k = 9 \times 10^9$ N m^2 C^{-2}.)

6 A point charge of 6.50 mC is suspended from a ceiling by an insulated rod. At what distance from the point charge will a small sphere of mass 10.0 g with a charge of –3.45 nC be located if it is suspended in air? (Use $k = 9 \times 10^9$ N m^2 C^{-2}.)

7 A charge of $+q$ is placed a distance r from another charge also of $+q$. A repulsive force of magnitude F is found to exist between them. Choose the correct answer from the options in bold to describe the changes, if any, that will occur to the force in the following:

a If one of the charges is doubled to $+2q$, the force will **halve**/**double**/**quadruple**/**quarter** and **repel**/**attract**.
b If both charges are doubled to $+2q$, the force will **halve**/**double**/**quadruple**/**quarter** and **repel**/**attract**.
c If one of the charges is changed to $-2q$, the force will **halve**/**double**/**quadruple**/**quarter** and **repel**/**attract**.
d If the distance between the charges is halved to $0.5r$, the force will **halve**/**double**/**quadruple**/**quarter** and **repel**/**attract**.

8 Calculate the repulsive force on each proton in a helium nucleus separated in a vacuum by a distance of 2.50 fm.
(Use $k = 9 \times 10^9$ N m^2 C^{-2}, 1 fm = 1×10^{-15} m and $q_p = +1.602 \times 10^{-19}$ C.)

9 Two point charges (30.0 cm apart in air) are charged by transferring electrons from one point to another. Calculate how many electrons must be transferred so that an attractive force of 1.0 N exists. Consider only the magnitude of q_e in your calculations.
(Use $k = 9 \times 10^9$ N m^2 C^{-2} and $q_e = -1.602 \times 10^{-19}$ C.)

2.3 The magnetic field

While naturally occurring magnets had been known for many centuries, by the early 19th century, there was still no scientifically proven way of creating an artificial magnet. In 1820, the Danish physicist Hans Christian Oersted (whose statue is pictured in Figure 2.3.1) developed a scientific explanation for the magnetic effect created by an electric current.

Oersted was a keen believer in the 'unity of nature', the concept that everything in the universe is somehow connected. He noticed that when he switched on a current from a **voltaic pile** (a simple early battery), a magnetic compass nearby moved. Intrigued by this observation, he carried out further experimentation, which demonstrated that it was the current from the voltaic pile that was affecting the compass movement. His experiments showed that the stronger the current, and the closer the compass was to it, the greater the observed effect. These observations led him to conclude that the electric current was creating a magnetic field. This connection between electric and magnetic fields is fundamental to society today.

FIGURE 2.3.1 In 1820, Hans Christian Oersted discovered the magnetic effect created by an electric current. Oersted is honoured by this statue in Oersted's Park, Copenhagen.

MAGNETISM

Before looking further into the connection between electric current and **magnetic fields**, it is necessary to review some fundamentals of magnetism.

The magnetic effect most people are familiar with is the attraction of iron or other **magnetic** materials to a magnet, as seen in Figure 2.3.2.

FIGURE 2.3.2 A bar magnet attracting drawing pins.

But, if you experiment with a magnet yourself, you will find that each end of a magnet behaves differently, particularly when interacting with another magnet. One end will be attracted while the other is repelled. Each end of a magnet is referred to as a **magnetic pole**.

Like magnetic poles repel each other; unlike magnetic poles attract each other.

EXTENSION

Dipoles

Try breaking a (cheap) magnet in half. All you get is two smaller magnets, each with its own north and south poles. No matter how many times you break the magnet and how small the pieces are, each will be a separate little magnet with two poles. Because magnets always have two poles, they are said to be *dipolar*.

Magnets are dipolar and a magnetic field is said to be a **dipole** field (see Figure 2.3.3). This is similar to electric charges where a positive and negative charge in close proximity to each other are said to form a dipole. A key difference is that you cannot have a single magnetic pole, whether it be a south pole or a north pole; however, charges *can* exist on their own as either a positive or negative charge.

FIGURE 2.3.3 Magnets are always dipolar.

A suspended magnet that is free to move will always orientate itself in a north–south direction. That's basically what the needle of a compass is—a freely suspended, small magnet. If allowed to swing vertically as well, then the magnet will tend to tilt vertically. The vertical direction (upwards/downwards) and the magnitude of the angle depend upon the direction of the magnet from either of the Earth's poles.

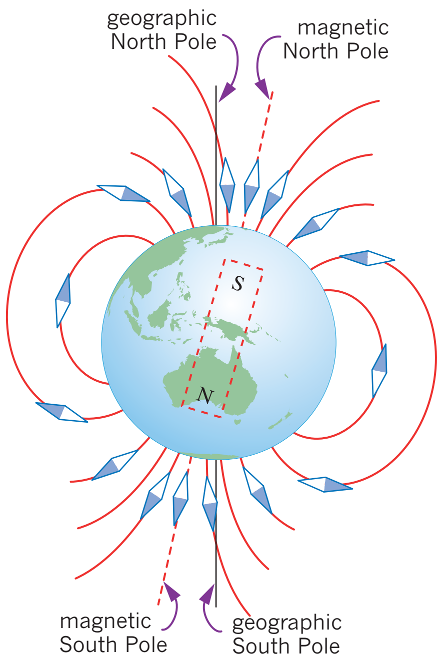


FIGURE 2.3.4 The Earth acts somewhat like a huge bar magnet. The south pole of this imaginary magnet is near the geographic North Pole and is the point to which the north pole of a compass appears to point.

As you can see in Figure 2.3.4, the Earth itself can be shown to have a giant magnetic field around it.

The names for the poles of a magnet derive from early observations of magnets orientating themselves with the Earth's geographic poles.

Initially, the end of the magnet pointing toward the Earth's geographic north was denoted the North **Pole**, and compasses are thus marked with this end as north. However, it is now known that the geographic North Pole and the magnetic North Pole are slightly apart in distance, and the same applies to the geographic South Pole and the magnetic South Pole.

PHYSICSFILE

'Flipping' poles

The Earth's magnetic poles are not static like their geographic counterparts. For many years, the magnetic North Pole had been measured as moving at around 9 km per year (see Figure 2.3.5). In recent years that has accelerated to an average of 52 km per year. Once every few hundred thousand years the magnetic poles actually flip in a phenomenon called 'geomagnetic reversal', so that a compass would point south instead of north. The Earth is well overdue for the next flip, and recent measurements have shown that the Earth's magnetic field is starting to weaken faster than in the past, so the magnetic poles may be getting close to a 'flip'. While past studies have suggested such a flip is not instantaneous—it would take many hundreds if not a few thousands of years—some more recent studies have suggested that it could happen over a significantly shorter time period.

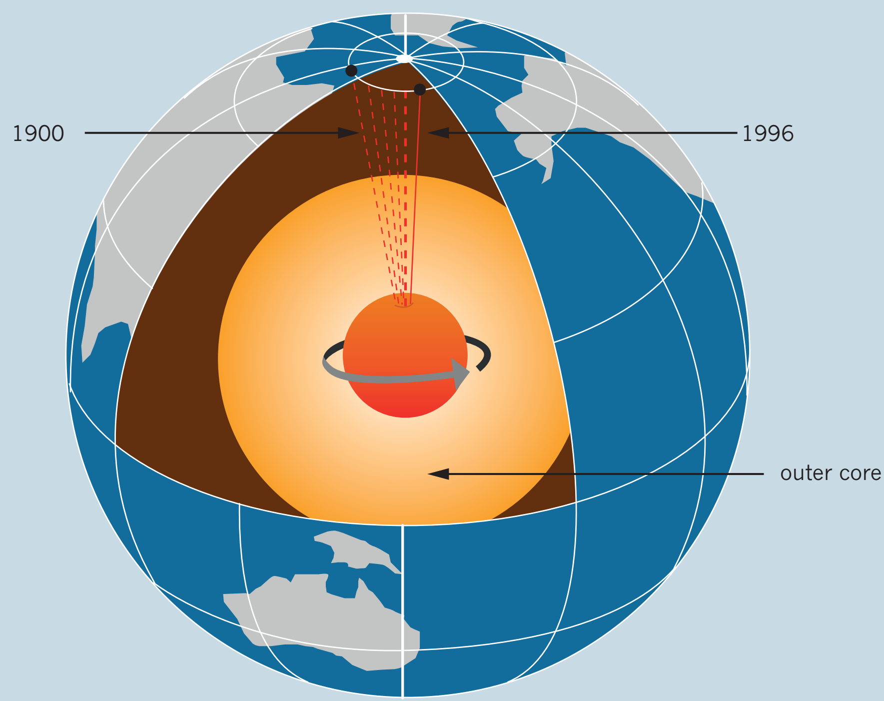


FIGURE 2.3.5 Diagram of Earth's interior and the movement of magnetic north from 1900 to 1996. The Earth's outer core is believed to be the source of the geomagnetic field.

MAGNETIC FIELD

In the earlier sections of this chapter, you saw that point charges and charged objects produce an electric field in the space that surrounds them. For this reason, charged bodies within the field will experience a force. The direction of the electric field is determined by the direction of that force.

Magnets also create fields. If you do a simple test like placing a pin near a magnet, you will observe that the pin will be pulled toward the magnet. This shows that the space around the magnet must therefore be affected by the magnet.

If you sprinkle iron filings on a piece of clear acetate that is held over a magnet, you will observe that the magnetic field will be clearly defined (see Figure 2.3.6). The iron filings will line up with the field, showing clear field lines running from one end of the magnet to the other.

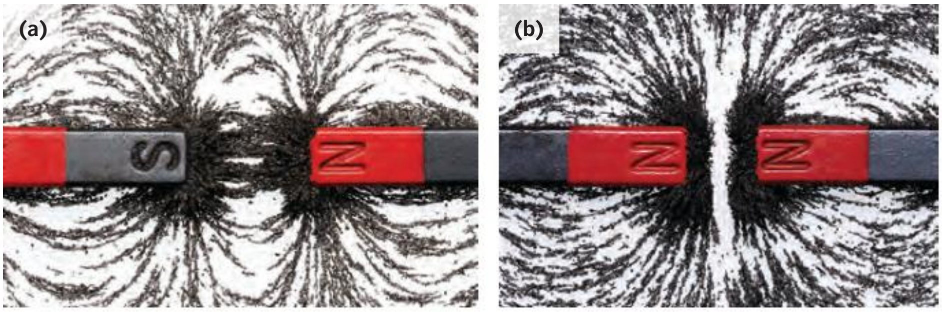


FIGURE 2.3.6 Iron fillings sprinkled around magnets (a) with unlike poles close together and (b) with like poles close together. The patterns in the fields show the attraction and repulsion between poles, respectively.

Vector field model for magnetic fields

The diagram in Figure 2.3.7 shows the magnetic field associated with a simple bar magnet. The magnetic field around the bar magnet can be defined in vector terms, specifying both direction and magnitude.

The direction of the magnetic field at any point is the direction that a compass would point if placed at that point—that is, towards the magnetic South Pole. This is also the direction of the force the magnetic field would exert on an (imaginary) single north pole.

Denser (closer) lines indicate a relatively stronger magnetic field. As the distance from the magnet increases, the magnetic field is spread over a greater area and its strength at any point decreases. The strength and direction associated with the magnetic field at any point signifies that it is a vector quantity. The strength, or vector magnitude, of the magnetic field at a particular point is denoted by B and has units of tesla (T).

The fields between magnets are dependent on whether like or unlike poles are close together, the distance the poles are apart and the relative strength of the magnetic field of each magnet. Iron filings or small plotting compasses can be used to visualise the field between and around the magnets, as shown in Figure 2.3.8.

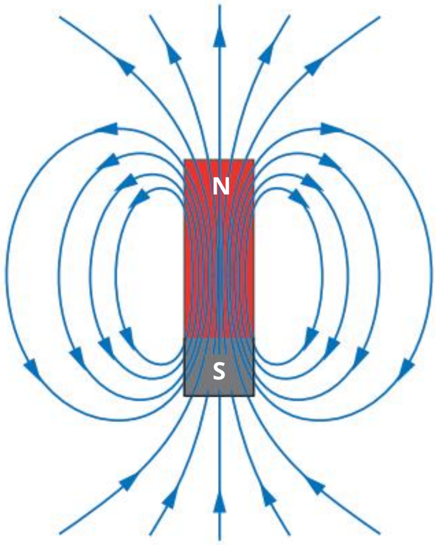


FIGURE 2.3.7 The field lines around and inside a bar magnet. The lines show the direction of the force on an (imaginary) single north pole.

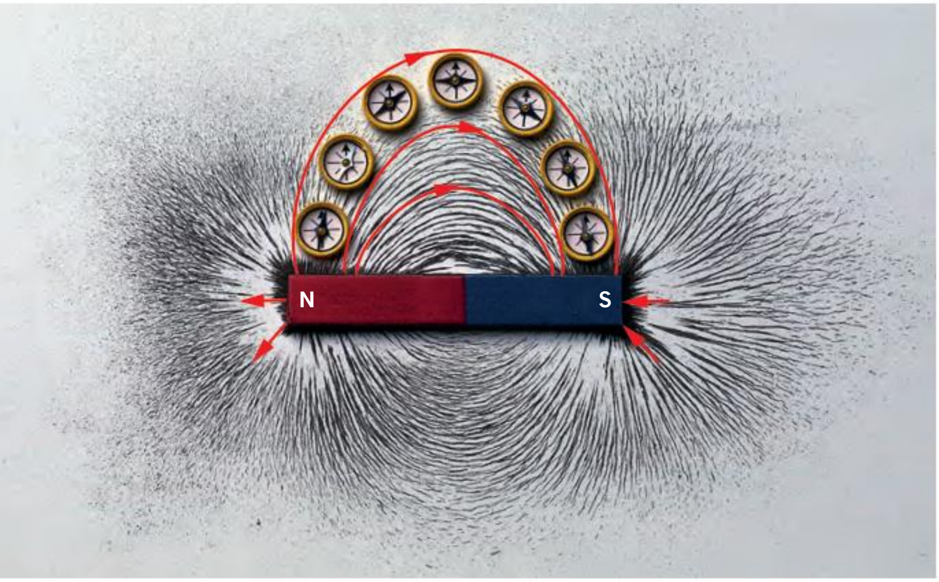


FIGURE 2.3.8 Plotting field lines around a bar magnet. Small plotting compasses are placed around the magnet. Field lines are drawn linking the direction each compass points in, creating field lines that run from the north pole to the south pole of the magnet.

Because the Earth has a giant magnetic field around it, you can also predict how compasses will orient themselves around the Earth—they will orient themselves along the magnetic field lines. In Figure 2.3.7, note the direction of the magnetic field close to either pole, where the magnetic field lines run almost vertically. Magnets placed near the Earth's magnetic poles will behave in the same way.

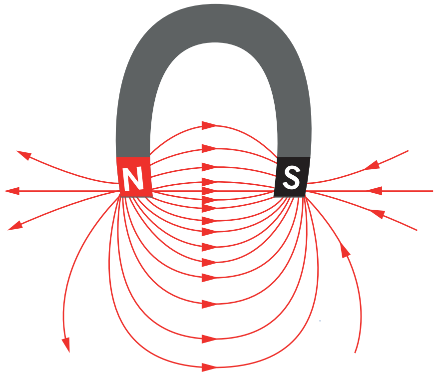


FIGURE 2.3.9 The horseshoe magnet has two unlike poles close to each other. This creates a very strong magnetic field.

Different shaped magnets produce different shaped fields. The diagram in Figure 2.3.9 shows the magnetic field plotted for a horseshoe magnet.

The resultant direction of the magnetic field at a particular point will be the vector addition of each individual magnetic field acting at that point.

When two magnets are placed close together, two situations may arise. If the poles are unlike as per Figure 2.3.10(b), then attraction will occur between them and a magnetic field will be created that extends between the two poles. On the other hand, if like poles are very near each other as per Figure 2.3.10(c), repulsion will occur. In this situation, there will be a neutral point between the two poles where there is no magnetic field.

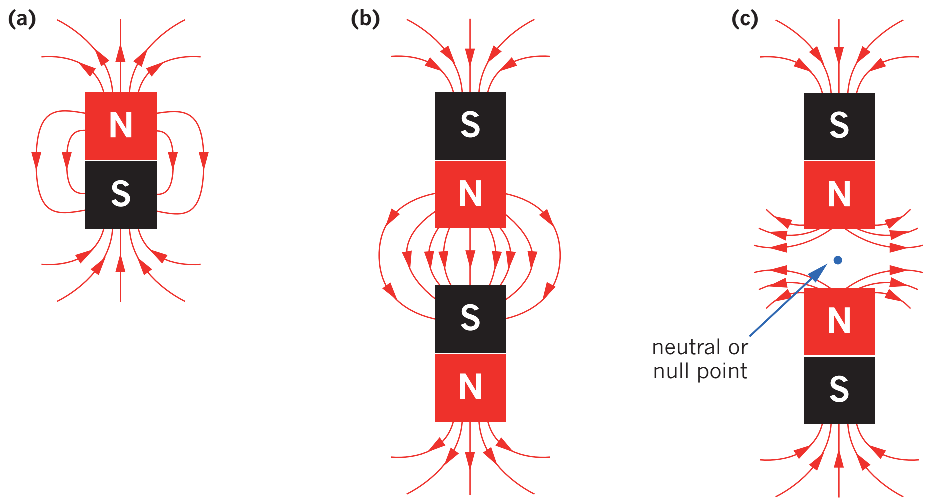


FIGURE 2.3.10 Magnetic field lines plotted for (a) a bar magnet, (b) opposite poles of magnets in close proximity and (c) like poles of magnets in close proximity.

As the bar magnets in Figure 2.3.10 have a fixed strength and position, the associated magnetic fields will be static. Varying the magnetic field strength, by changing the magnets or varying the relative position of the magnets, would produce a changing magnetic field.

MAGNETIC FIELDS AND CURRENT-CARRYING WIRES

In the introduction to this section the connection between electric current and magnetic fields was noted. Oersted found that when he switched on the current from a voltaic pile, a nearby magnetic compass would move. It's believed that the Earth's magnetic field is created by a similar effect—circulating electric currents in the Earth's molten metallic core.

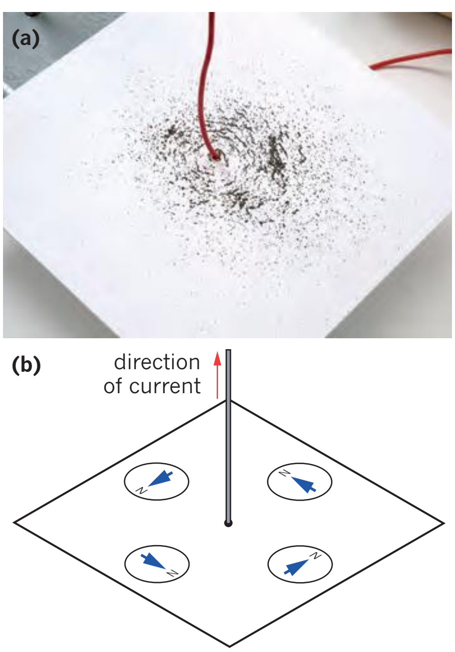


FIGURE 2.3.11 (a) The magnetic field around a current-carrying wire. The iron filings align with the field to show the circular nature of the magnetic field. (b) Small compasses will indicate the direction of the field.

A circular magnetic field is created around a current-carrying wire. This can be seen in Figure 2.3.11. A compass aligns itself at a tangent to the concentric circles around the wire (i.e. the magnetic field). The stronger the current and the closer the compass is to the wire, the greater the effect.

The magnetic field is perpendicular to the current-carrying wire and the direction of the field will depend upon the current direction. There's a simple and easy way to determine the direction of the magnetic field, which is commonly referred to as the **right-hand grip rule**.

Grasp the conducting wire with your right hand with your thumb pointing in the direction of the conventional electric current, *I* (positive to negative). Curl your fingers around the wire. The magnetic field will be perpendicular to the wire and in the direction your fingers are pointing, as shown in Figure 2.3.12 on the following page.

Worked example 2.3.1

DIRECTION OF THE MAGNETIC FIELD

A current-carrying wire runs horizontally across a table. The conventional current direction, *I*, is running from left to right. What is the direction of the magnetic field created by the current?

Thinking	Working
Recall that the right-hand grip rule indicates the direction of the magnetic field.	Hold your hand with your fingers aligned as if gripping the wire. Point your thumb to the right in the direction of the current flow. 
Describe the direction of the field in relation to the reference object or wire in simple terms, so that the description can be readily understood by a reasonable reader.	The magnetic field direction is perpendicular to the wire and runs from up the back of the wire, over the top towards the front of the wire.

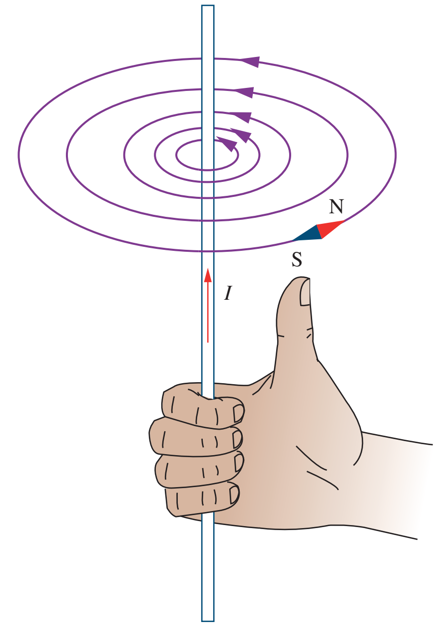


FIGURE 2.3.12 The right-hand grip rule can be used to find the direction of the magnetic field around a current-carrying wire, when the direction of the conventional current, *I*, is known.

Worked example: Try yourself 2.3.1

DIRECTION OF THE MAGNETIC FIELD

A current-carrying wire runs along the length of a table. The conventional current direction, *I*, is running toward an observer standing at the near end. What is the direction of the magnetic field created by the current as seen by the observer?

Magnetic fields between parallel wires

Two current carrying wires arranged parallel to each other will each have their own magnetic field. The direction of the magnetic field around each wire is given by the right-hand grip rule. If the two wires are brought closer together, their associated magnetic fields will interact, just as any two regular magnets would interact. The interaction could result in either an attraction or repulsion of the wires, depending on the direction of the magnetic fields between them (see Figure 2.3.13). When the magnetic fields are in the opposite directions, this represents unlike poles, and so the wires attract. When the magnetic fields are in the same direction, the wires repel.

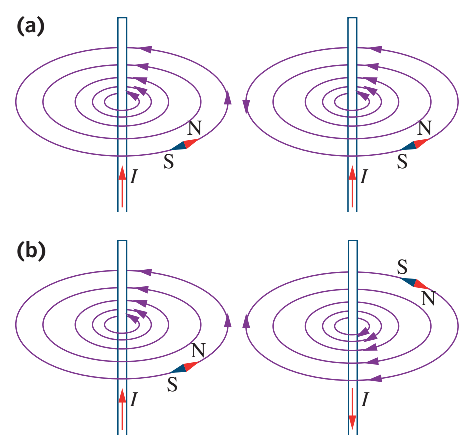


FIGURE 2.3.13 (a) Two current-carrying wires attract when current runs through them in the same direction. This is because the magnetic fields between the wires are in opposite directions. (b) Two current-carrying wires repel when the current passes through them in opposite directions. This is because the magnetic fields are in the same direction.

3D FIELDS

Field lines can also be drawn for more-complex, 3D fields such as that around the Earth or those around current-carrying loops and coils. Even in more-complex fields, the right-hand grip rule is still applicable, as you can see in figures 2.3.14 and 2.3.15.

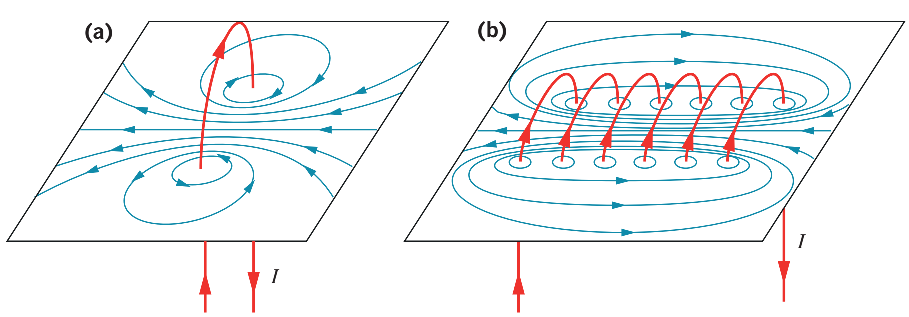


FIGURE 2.3.14 The magnetic field lines around (a) a single current loop and (b) a series of loops. The blue arrows indicate the direction of the magnetic field. The more concentrated the lines are inside the loops, the stronger the magnetic field is in this region.

The direction of a magnetic field can be shown with a simple arrow on a field line when the field is travelling within the plane of a page (as shown in Figure 2.3.14), or a simple 3D depiction can be used as in Figure 2.3.15.

Figure 2.3.15 shows a 3D representation of the magnetic field around a loop of wire. This same loop can also be represented in two dimensions using the following conventions.

When a field is running directly into or out of the plane of a page, dots are used to show a field coming out of the page and crosses are used to depict a field running directly into the page. This convention was adopted from the idea of viewing an arrow. The dot is the point of the arrow coming toward you, and the cross represents the tail feathers as the arrow travels away.

Figure 2.3.16 shows the 2D representation of the same magnetic field around a simple loop of wire that was shown in Figure 2.3.15.

The strength of a field is depicted by varying the density of the lines or dots and crosses. Showing lines coming closer together indicates a strengthening of a field, less density indicates a weaker field. More densely placed dots or crosses can also show a stronger area of the field. This would be referred to as a non-uniform magnetic field.

As the magnetic fields associated with current-carrying coils are dependent upon the size of the current, the associated field may also be changing over time, either in magnitude, or, if the current is reversed, in direction.

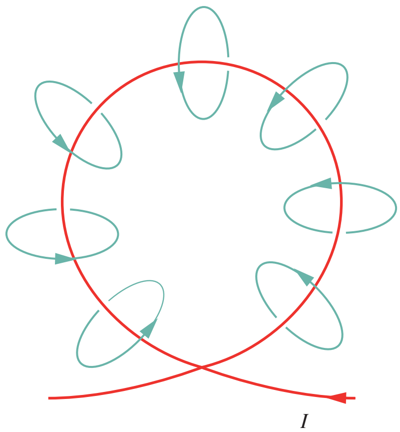


FIGURE 2.3.15 A 3D representation of the magnetic field around a loop of wire in the plane of the page. The blue arrows show the direction of the magnetic field. Notice that the magnetic field is a circular shape, with no field lines crossing.

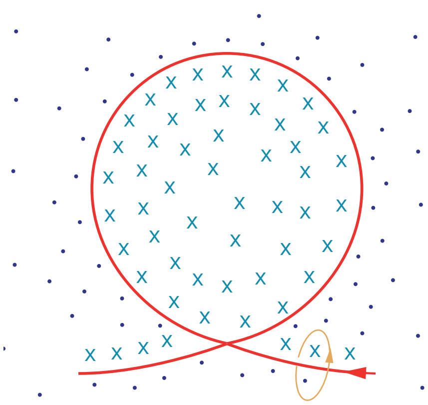


FIGURE 2.3.16 A 2D representation of the same current-carrying loop depicted in Figure 2.3.15. Areas where the magnetic field is stronger are shown with a greater density of dots and crosses.

The magnetic field around a solenoid

If many loops are placed side by side, their fields all add together and there is a much stronger effect. This can easily be achieved by winding many turns of wire into a coil termed a **solenoid**. The field around the solenoid is like the field around a normal bar magnet. The direction of the overall magnetic field can be determined by considering the field around each loop and, in turn, the field around the current-carrying wire making up the loop. The direction of the field of the solenoid depends on the direction of the current in the wire making up the solenoid.

This is explained in the diagram in Figure 2.3.17.

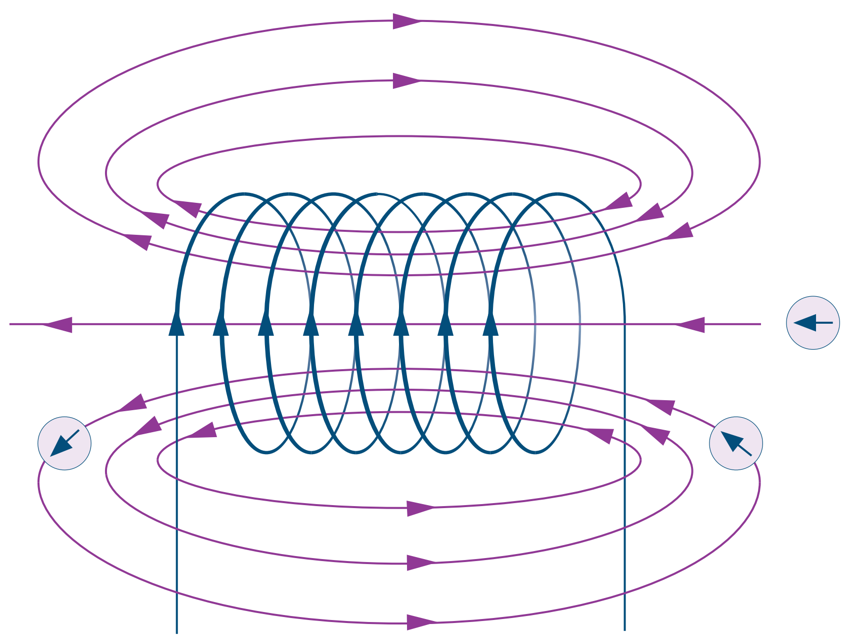


FIGURE 2.3.17 This solenoid has an effective 'north' end at the left and a 'south' end at the right. The compass points in the direction of the field lines.

A simple way to remember which end of a solenoid is which pole is to write the letters S and N and put arrows on them as shown in Figure 2.3.17. The arrows indicate the direction of the current as seen from that end. Try each of S and N at each end to determine which pole 'fits' at which end.

CREATING AN ELECTROMAGNET

The earliest magnets were all naturally occurring. If you wanted a magnet, you needed to find one. They were regarded largely as curiosities. Hans Christian Oersted's discoveries made it possible to manufacture magnets, making the widescale use of magnets possible.

An **electromagnet**, as the name infers, runs on electricity. It works because an electric current produces a magnetic field around a current-carrying wire. If the conductor is looped into a series of coils to make a solenoid, then the magnetic field can be concentrated within the coils. The more coils, the stronger the magnetic field and, therefore, the stronger the electromagnet.

The magnetic field can be strengthened further by wrapping the coils around a core. Normally, the atoms in materials like iron point in random directions and the individual magnetic fields tend to cancel each other out. However, the magnetic field produced by coils wrapped around an iron core can force the atoms within the core to point in one direction. Their individual magnetic fields add together, creating a stronger magnetic field.

The strength of an electromagnet can also be changed by varying the amount of electric current that flows through it.

The direction of the current creates poles in the electromagnet. The poles of an electromagnet can be reversed by reversing the direction of the electric current.

Today, electromagnets are used directly to lift heavy objects (see Figure 12.3.18), as switches and relays, and as a way of creating new permanent magnets by aligning the atoms within magnetic materials.

FIGURE 2.3.18 A large electromagnet being used to lift waste iron and steel at a scrapyard. Valuable metals such as these are separated and then recycled.

2.3 Review

SUMMARY

- Like magnetic poles repel, and unlike magnetic poles attract.
- Magnetic poles exist only as dipoles, having both north and south poles. A single magnetic pole (monopole) is not known to exist.
- The direction of a magnetic field at a particular point is the same as that of the force on a (imaginary) single north pole.
- The resultant direction of interacting magnetic fields at any particular point will be the vector addition of each individual magnetic field acting at that point.
- The Earth has a dipolar magnetic field that acts as a huge bar magnet, with the south end near the geographic North Pole.
- A magnetic field associated with a constant magnetic field is static. Where the magnetic field is changing, such as that associated with an alternating current direction, the magnetic field will also be changing.
- A uniform distribution of field lines represents a uniform magnetic field. A non-uniform field, such as that around a non-circular coil, is shown by variations in the separation of the field lines.
- An electrical current produces a magnetic field that is circular around a current-carrying conductor. The direction of the field is given by using the *right-hand grip rule* when considering the direction of the conventional current.
- More complex fields can be determined by applying the right-hand grip rule to the loops or coils making up the current carrying conductor in a solenoid.

KEY QUESTIONS

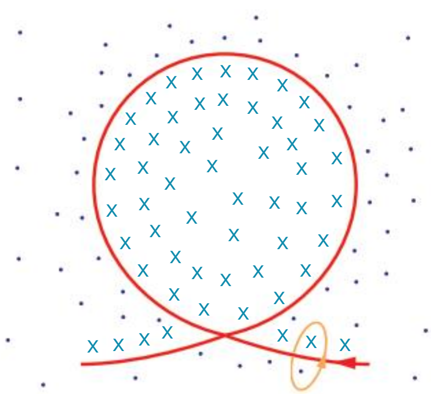
1 Repeatedly cutting a magnet in half always produces magnets with two opposite poles. From this information, which of the following can be deduced in relation to the poles of a magnet?

A Magnets are easily sliced in half.
B All magnets are dipolar.
C When the magnets are cut, the poles are split in half.
D All split magnets are monopolar.

2 A magnet is suspended on a thin wire at its midpoint so that it is free to swing. In which direction, approximately, will the north pole of the magnet point?

A Earth's geographic north
B Earth's geographic south
C Earth's equator
D the sky due to Earth's gravity

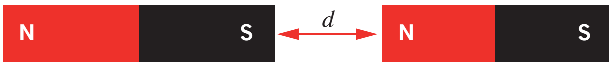
3 The field around a particular current-carrying loop shows a variation in the strength of its magnetic field, as depicted below. The current in the loop itself is being switched on and off but is constant in direction and size.



Which of the following best describes the resulting magnetic field of the loop?

A a reversed magnetic field
B a static magnetic field
C a non-uniform magnetic field
D a uniform magnetic field

4 The following diagram shows two bar magnets separated by a distance *d*. At this separation, the magnitude of the magnetic force between the poles is equal to *F*. Which of the following is true if the distance, *d*, is increased?



A An attractive force greater than *F* will exist between the poles.

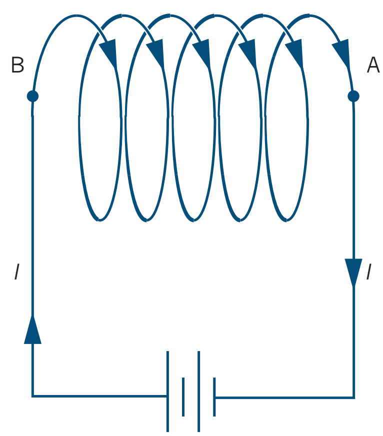
B A repulsive force greater than *F* will exist between the poles.

C An attractive force less than *F* will exist between the poles.

D A repulsive force less than *F* will exist between the poles.

5 A current-carrying wire runs horizontally across a table. The conventional current direction, *I*, is running from right to left. Draw a diagram showing the direction of the magnetic field around the wire.

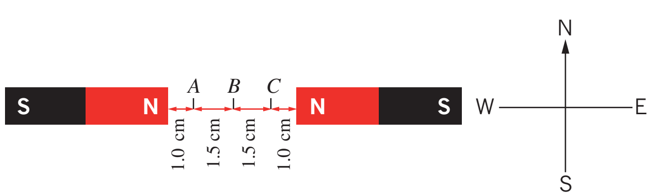
6 The following diagram shows a current-carrying solenoid.



Which end (A or B) represents the north pole of this solenoid?

The following information applies to questions 7–9.

Two strong bar magnets which produce magnetic fields of equal strength are arranged as shown.

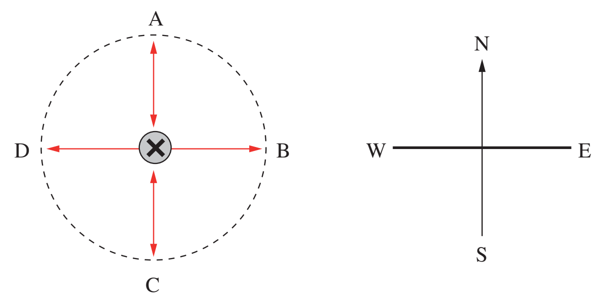


7 Ignoring the magnetic field of the Earth, what is the approximate direction of the resulting magnetic field at point A?

8 Ignoring the magnetic field of the Earth, what is the approximate direction of the resulting magnetic field at point C?

9 Ignoring the magnetic field of the Earth, what is the magnitude of the resulting magnetic field at point B?

10 The figure below shows a cross-sectional view of a long, straight, current-carrying conductor, with its axis perpendicular to the plane of the page. The conductor carries an electric current into the page.



a What is the direction of the magnetic field produced by this conductor at each of the points A, B, C and D?

b The direction of the current in the conductor is now changed so that it is carried out of the page. What is the direction of the magnetic field produced by this conductor at the four points A, B, C and D?

2.4 Forces on charged objects due to magnetic fields

An electric current is a flow of electric charges. These may be electrons in a metal wire, electrons and mercury ions in a fluorescent tube or cations and anions in an electrolytic cell. The nature of the flowing charge that makes up the current does not matter. A magnetic field is produced around the flow of charge, and a force is experienced within this field (see Figure 2.4.1). In each case, it is the total rate of flow of charge, i.e. the current, which determines the field produced or the magnitude of the force.

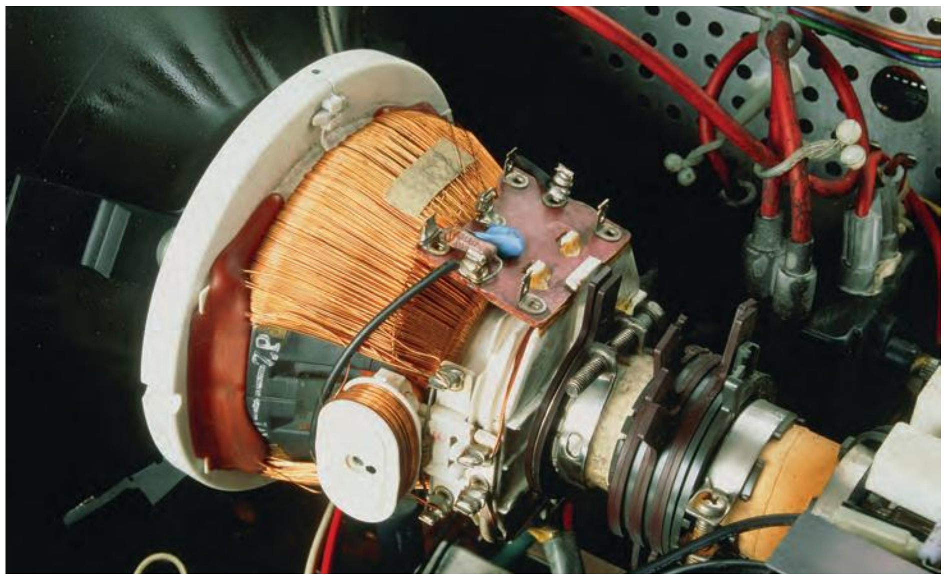


FIGURE 2.4.1 Electrons rushing down the length of a CRT (cathode ray tube) were the basis upon which old-style television sets worked. The electrons were deflected by the magnetic force they experienced as they passed through the 'yoke'—coils of copper wire at the back of the tube creating a strong variable magnetic field.

MAGNETIC FORCE ON CHARGED PARTICLES

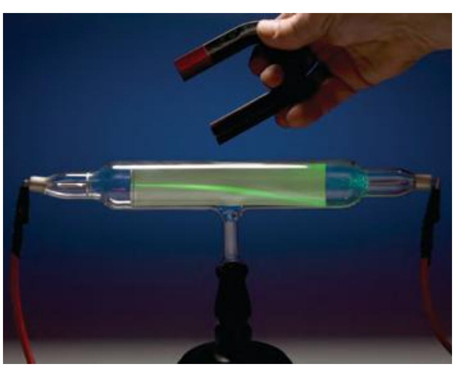


FIGURE 2.4.2 The electron beam of a cathode ray tube being deflected by a magnet.

The principle behind a **cathode ray tube** (CRT) is that a charged particle moving within a magnetic field will experience a force. In Figure 2.4.2, a beam of electrons in a CRT is experiencing a force due to a magnetic field. The force causes the beam of electrons to bend. The magnitude of the force is proportional to the strength of the magnetic field, *B*, the component of the velocity of the charge that is perpendicular (at right angles) to the magnetic field and the charge on the particle; that is:

When v and B are perpendicular:

$$F = qvB$$

where F is the force in newtons (N)

q is the electric charge on the particle in coulombs (C)

v is the component of the instantaneous velocity of the particle that is perpendicular to the magnetic field (m s^{-1})

B is the strength of the magnetic field (T)

This force is referred to as the **Lorentz force**. The force is at a maximum when the charged particle is moving at right angles to the field. There is no force acting when the charged particles are travelling parallel to the magnetic field.

Determining the direction of the force

The simple **mnemonic** shown in Figure 2.4.3 can be used to determine the direction of the force on a charged particle moving in a magnetic field. Using your right hand, with fingers outstretched and flat, point the thumb toward the direction that a positive charge is moving and the outstretched fingers in the direction of the magnetic field. The direction of the resulting force on the charge is the direction in which your palm is pointing. The force on a negatively charged particle will therefore be in the *opposite* direction to that on a positively charged particle.

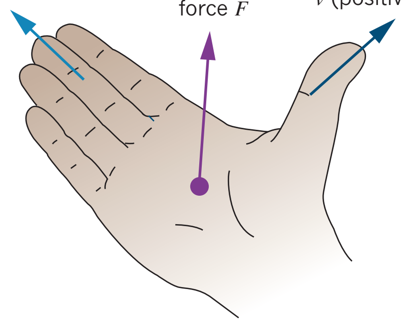


FIGURE 2.4.3 The right-hand rule: Point the thumb of the right hand in the direction of the movement of *a positive charge* (conventional current direction) and the fingers in the direction of the magnetic field. The force on the charge will point out from the palm.

Worked example 2.4.1

MAGNITUDE OF FORCE ON A POSITIVELY CHARGED PARTICLE

A single, positively charged particle with a charge of $+1.6 \times 10^{-19}$ C travels at a velocity of 10 m s^{-1} perpendicular to a magnetic field, B, of strength 4.0×10^{-5} T. What is the magnitude of the force the particle will experience from the magnetic field?

Thinking	Working
Check the direction of the velocity and determine whether a force will apply. Forces only apply on the component of the velocity perpendicular to the magnetic field.	The particle is moving perpendicular to the field. A force will apply, and so $F = qvB$.
Establish which quantities are known and which ones are required.	$F = ?$ $q = +1.6 \times 10^{-19}$ C $v = 10$ m s^{-1} $B = 4.0 \times 10^{-5}$ T
Substitute values into the force equation.	$F = qvB$ $= 1.6 \times 10^{-19} \times 10 \times 4.0 \times 10^{-5}$
Express the final answer in an appropriate form. Note that only magnitude has been requested so do not include direction.	$F = 6.4 \times 10^{-23}$ N

Worked example: Try yourself 2.4.1

MAGNITUDE OF FORCE ON A POSITIVELY CHARGED PARTICLE

A single, positively charged particle with a charge of $+1.6 \times 10^{-19}$ C travels at a velocity of 50 m s^{-1} perpendicular to a magnetic field, B, of strength 6.0×10^{-5} T. What is the magnitude of the force the particle will experience from the magnetic field?

PHYSICSFILE

The tesla

The unit for the strength of a magnetic field, B, was given the name tesla (T) in honour of Nikola Tesla. Nikola Tesla (1856–1943) was the first person to advocate the use of alternating current (AC) generators for use in town power-supply systems. He was also a prolific inventor of electrical machines of all sorts, including the Tesla coil, a source of high-frequency, high-voltage electricity.

A magnetic field of 1 T is a very strong field. For this reason, a number of smaller units, especially the millitesla (mT), 10^{-3} T, and microtesla (μT), 10^{-6} T, are in common use. The table below shows the strength of some magnets for comparison.

Type of magnet	Strength of magnetic field
very strong electromagnets and ‘super magnets’	1 to 20 T
Alnico and ferrite magnets	10^{-2} to 1 T
Earth’s surface	5×10^{-5} T

TABLE 2.4.1 Comparison of magnet strengths.

EXTENSION

Objects moving at an angle to the magnetic field

The force experienced by a charge moving in a magnetic field is a vector quantity. The original expression noted above applies only to that component of the velocity of the charge perpendicular to the magnetic field. To find the force acting on an object moving at an angle θ to the magnetic field, use:

$F = qvB \sin \theta$

A charged particle travelling at a steady speed in a magnetic field experiences this force at an angle to its path and will be diverted.

This is the theory behind CRT screens. As the direction of the charged particle changes, so does the angle of the force acting on it. In a very large magnetic field the charged particles will move in a circular path. Mass spectrometers and particle accelerators both work on this principle.

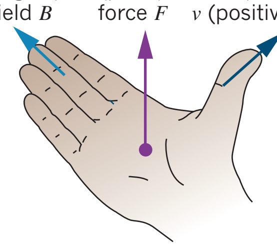
When high-energy particles in the **solar wind** from the Sun meet the Earth's magnetic field, they also experience this type of force. As the particles approach the Earth, they encounter the magnetic field and are deflected in such a way that they spiral towards the poles, losing much of their energy and creating the auroras (the southern aurora, or aurora australis, and the northern aurora, or aurora borealis, as shown in Figure 2.4.4).

FIGURE 2.4.4 Charged particles from the Sun or deep space are trapped by the Earth's magnetic field, causing them to spiral towards the poles. As they do this, they lose energy and create the auroras.

Worked example 2.4.2

DIRECTION OF FORCE ON A NEGATIVELY CHARGED PARTICLE

A single, negatively charged particle with a charge of -1.6×10^{-19} C is travelling horizontally out of a computer screen and perpendicular to a magnetic field, *B*, that runs horizontally from left to right across the screen. In what direction will the force experienced by the charge act?

Thinking	Working
 The right-hand rule is used to determine the direction of the force on a positively charged particle.	Align your hand so that your fingers are pointing in the direction of the magnetic field, i.e. left to right and horizontal. If the negatively charged particle is travelling out of the screen, a positively charged particle would be moving in the opposite direction. Align your thumb so it is pointing into the screen, in the direction that a positive charge would travel. Your palm should be facing downwards. That is the direction of the force applied by the magnetic field on the negative charge out of the screen.

Worked example: Try yourself 2.4.2

DIRECTION OF FORCE ON A NEGATIVELY CHARGED PARTICLE

A single, negatively charged particle with a charge of -1.6×10^{-19} C is travelling horizontally from left to right across a computer screen and perpendicular to a magnetic field, *B*, that runs vertically down the screen. In what direction will the force experienced by the charge act?

THE FORCE ON A CURRENT-CARRYING CONDUCTOR

Since a conducting wire is essentially a stream of charged particles flowing in one direction, it is not hard to imagine that a conductor carrying a stream of charges within a magnetic field will also experience a force. This is the theory behind the operation of electric motors that will be explained in the chapter 'Applications of fields'.

The current in a conductor is dependent on the rate at which charges are moving through the conductor; that is:

$$I = \frac{Q}{t}$$

where I is the current (A)

Q is the total charge (C)

t is the time taken (s).

For a 1 m length of conductor, the velocity of the charges through the conductor is:

$$v = \frac{s}{t} = \frac{1}{t}$$

And hence

$$I = \frac{Q}{t} = Q \times \frac{1}{t} = Qv$$

As $F = qvB$ for a single charge, q, moving perpendicular to a magnetic field, then:

$F = IB$ for a one metre conductor,

and for a conductor of any length, l, $F = IlB$

and for a conductor made up of n loops or conductors of length l:

$$F = nIlB$$

where F is the force on the conductor perpendicular to the magnetic field in newtons (N)

n is the number of loops or conductors

I is the current in the conductor in amperes (A)

l is the length of the conductor in metres (m)

B is the strength of the magnetic field in tesla (T)

Just as for a single charge moving in a magnetic field, the force on the conductor is at a maximum when the conductor is at right angles to the field. The force is zero when the conductor is parallel to the magnetic field. The right-hand rule is used to determine the direction of the force.

EXTENSION

Conductors at an angle to a magnetic field

The force experienced by a current-carrying conductor is a vector quantity. The expression noted above applies only to that component of the conductor perpendicular to the magnetic field. To find the force acting on any conductor, or part of a conductor, moving at an angle θ to the magnetic field, use the equation:

$$F = nIlB \sin \theta$$

This is particularly relevant when applied to practical electric motors.

PHYSICS IN ACTION

The current balance

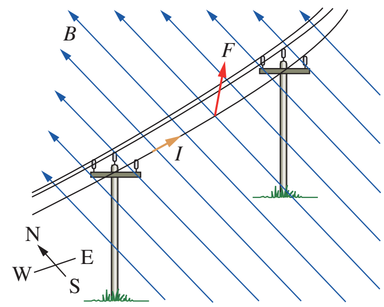
A current balance can be used to determine the force on a conductor in a magnetic field, as shown in Figure 2.4.5.

FIGURE 2.4.5 A current balance is used to measure the interaction between an electric conductor and a magnetic field. The relationship between force, current and conductor length can be shown.

Worked example 2.4.3

MAGNITUDE OF THE FORCE ON A CURRENT-CARRYING WIRE

Determine the magnitude of the force due to the Earth's magnetic field that acts on a suspended power line running east–west near the equator at the moment it carries a current of 100 A from west to east. Assume that the strength of the Earth's magnetic field at this point is 5.0×10^{-5} T.

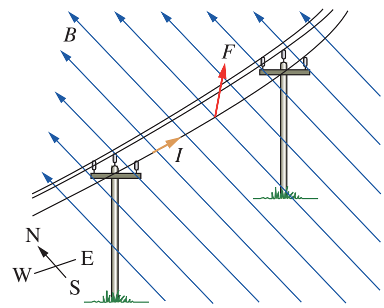


Thinking	Working
Check direction of the conductor and determine whether a force will apply. Forces only apply to the component of the wire perpendicular to the magnetic field.	As the current is running west–east and the Earth's magnetic field runs south–north, the current and the field are at right angles and a force will exist.
Establish what quantities are known and what are required. Since the length of the power line hasn't been supplied, consider the force per unit length (i.e. 1 m).	$F = ?$ $n = 1$ $I = 100$ A $l = 1.0$ m $B = 5.0 \times 10^{-5}$ T
Substitute values into the force equation and simplify.	$F = nIlB$ $= 1 \times 100 \times 1.0 \times 5.0 \times 10^{-5}$ N $= 5.0 \times 10^{-3}$ N
Express the final answer in an appropriate form with a suitable number of significant figures. Note that only magnitude has been requested, so do not include direction.	$F = 5.0 \times 10^{-3}$ N per metre of power line

Worked example: Try yourself 2.4.3

MAGNITUDE OF THE FORCE ON A CURRENT-CARRYING WIRE

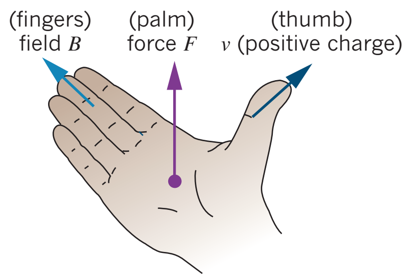
Determine the magnitude of the force due to the Earth's magnetic field that acts on a suspended power line running east–west near the equator at the moment it carries a current of 50 A from west to east. Assume that the strength of the Earth's magnetic field at this point is 5.0×10^{-5} T.



Worked example 2.4.4

DIRECTION OF THE FORCE ON A CURRENT-CARRYING WIRE

A current balance is used to measure the force from a magnetic field on a wire of length 5.0 cm running perpendicular to the magnetic field. The conventional current direction in the wire is from left to right. The magnetic field can be considered to be running into the page. What is the direction of the force on the wire?

Thinking	Working
 The right-hand rule is used to determine the direction of the force.	Align your hand so that your fingers are pointing in the direction of the magnetic field, i.e. into the page. Align your thumb so it is pointing right, in the direction of the current. Your palm should be facing upwards. That is the direction of the force applied by the magnetic field on the wire.
State the direction in terms of the other directions included in the question. Make the answer as clear as possible to avoid any misunderstanding.	The force on the wire is acting vertically upwards.

Worked example: Try yourself 2.4.4

DIRECTION OF THE FORCE ON A CURRENT-CARRYING WIRE

A current balance is used to measure the force from a magnetic field on a wire of length 5.0 cm running perpendicular to the magnetic field. The conventional current direction in the wire is from left to right. The magnetic field can be considered to be running out of the page. What is the direction of the force on the wire?

Worked example 2.4.5

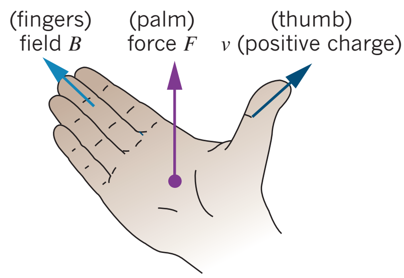
FORCE AND DIRECTION ON A CURRENT-CARRYING WIRE

The Amundsen–Scott South Pole Station sits at a point that can be considered to be at the Earth's southern magnetic pole (which behaves like the north pole of a magnet).

Assuming the strength of the Earth's magnetic field at this point is 5.0×10^{-5} T, determine the magnitude and direction of the magnetic force on the following:

a A 2.0 m length of wire carrying a conventional current of 10.0 A vertically up the exterior wall of one of the buildings.

Thinking	Working
Forces only apply to the components of the wire running perpendicular to the magnetic field. The direction of the magnetic field at the southern magnetic pole will be almost vertically upwards.	The section of the wire running up the wall of the building will be parallel to the magnetic field, *B*. Hence, no force will apply.
State your answer. A numeric value is required. Since there is no force, it is not necessary to state a direction.	$F = 0$ N

b A 2.0 m length of wire carrying a conventional current of 10.0 A running horizontally right to left across the exterior of one of the buildings.	
Forces only apply to the components of the wire running perpendicular to the magnetic field. The direction of the magnetic field at the southern magnetic pole will be almost vertically upwards (that is, out of the ground).	The section of the wire running horizontally through the building will be perpendicular to the magnetic field, *B*. A force *F* with a strength equivalent to *nIlB* will apply.
Identify the known quantities.	$F = ?$ $n = 1$ $I = 10.0$ A $l = 2.0$ m $B = 5.0 \times 10^{-5}$ T
Substitute into the appropriate equation and simplify.	$F = nIlB$ $= 1 \times 10.0 \times 2.0 \times 5.0 \times 10^{-5}$ $= 1.00 \times 10^{-3}$ N
 The direction of the magnetic force is also required to fully specify the vector quantity. Determine the direction of the magnetic force using the right-hand rule.	Align your hand so that your fingers are pointing in the direction of the magnetic field, i.e. vertically up. Align your thumb so it is pointing left, in the direction of the current. Your palm should be facing inwards (towards the building). That is the direction of the force applied by the magnetic field on the wire.
State the magnetic force in an appropriate form with a suitable number of significant figures. Include the direction to fully specify the vector quantity.	$F = 1.0 \times 10^{-3}$ N inwards

Worked example: Try yourself 2.4.5

FORCE AND DIRECTION ON A CURRENT-CARRYING WIRE

Santa's house sits at a point that can be considered the Earth's magnetic North Pole (which behaves like the south pole of a magnet).

Assuming the strength of the Earth's magnetic field at this point is 5.0×10^{-5} T, calculate the magnetic force and its direction on the following:

a a 2.0 m length of wire carrying a conventional current of 10.0 A vertically up the outside wall of Santa's house.

b a 2.0 m length of wire carrying a conventional current of 10.0 A running horizontally right to left across the outside of Santa's house.

2.4 Review

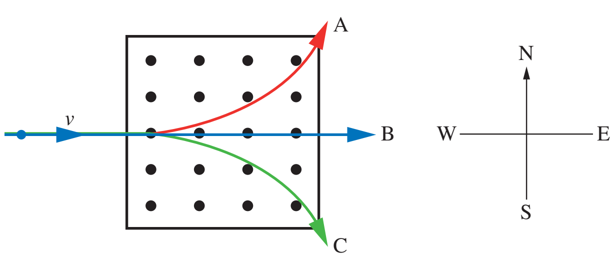
SUMMARY

- The magnitude of the force on a charged object within a magnetic field is proportional to the strength of the magnetic field, B, the component of the velocity of the charge that is perpendicular (at right angles) to the magnetic field, and the charge on the particle, i.e. $F = qvB$.
- This force is referred to as the Lorentz force.
- The force is at a maximum when the charged particle is moving at right angles to the magnetic field.
- The force is zero when the charged particle is travelling parallel to the magnetic field.
- The right-hand rule is used to determine the direction of the force on a positive charge moving in a magnetic field, B. The direction of the force on a negatively charged particle is in the opposite direction.
- The magnetic force on a current-carrying wire within a magnetic field is
  $F = nIlB$
- The direction of the force is given by the right-hand rule where the force travels out of the palm of the hand, once the thumb and fingers are orientated in the direction of the (conventional) current and magnetic field, respectively.

KEY QUESTIONS

1 A single, positively charged particle with a charge of $+1.6 \times 10^{-19}$ C is travelling into a computer screen and perpendicular to a magnetic field, B, that runs horizontally from left to right across the screen.
In what direction will the force experienced by the charge act?
A left to right
B right to left
C vertically up
D vertically down

2 The following diagram shows a particle, with initial velocity v, about to enter a uniform magnetic field, B, directed out of the page.

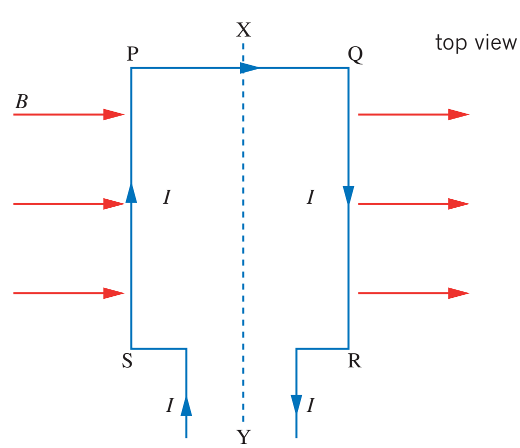


a If the charge on this particle is positive, what is the direction of the force on this particle just as it enters the field?
b Which path will this particle follow?
c Does the kinetic energy of the particle increase decrease or remain constant?
d If this particle were negatively charged, what path would it follow?
e What kind of particle could follow path B?

3 A single, positively charged particle with a charge of $+1.6 \times 10^{-19}$ C travels at a velocity of 0.5 m s^{-1} from left to right perpendicular to a magnetic field, B, of strength 2.0×10^{-5} T, running vertically downwards. What is the magnitude of the force that the particle will experience from the magnetic field?
A 1.6×10^{-5} N
B 3.2×10^{-5} N
C 1.6×10^{-19} N
D 1.6×10^{-24} N

4 A single, negatively charged particle with a charge of -1.6×10^{-19} travels at a velocity of 1.0 m s^{-1} from right to left parallel to a magnetic field, B, of strength 3.0×10^{-5} T.
What is the magnitude of the force the particle will experience from the magnetic field?

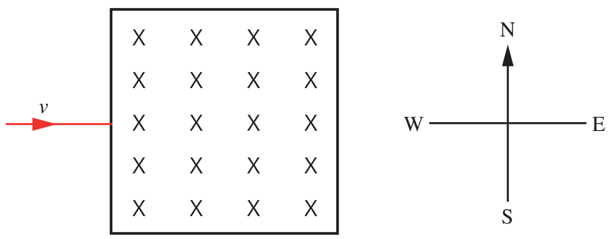
5 A rectangular loop of wire is carrying a current, I, in a magnetic field, B, as shown below. What is the direction of the force on the length of wire marked PQ?



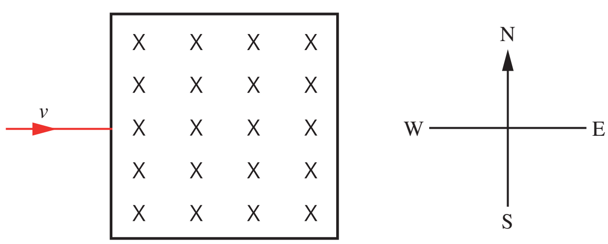
2.4 Review *continued*

6 An east–west power line of length 100 m is suspended between two towers. Assume that the strength of the magnetic field of the Earth in this region is 5.0×10^{-5} T. Calculate the magnetic force (including direction) on this power line at the moment it carries a current of 80 A from west to east.

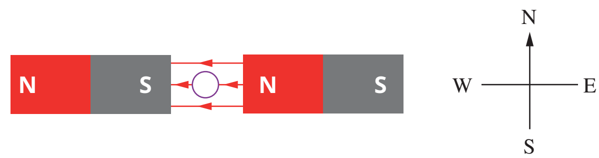
7 An electron with a charge of -1.6×10^{-19} C is moving eastwards into a magnetic field of strength 1.5×10^{-5} T acting into the page, as shown below. If the magnitude of the initial velocity is 2 m s^{-1}, what is the magnitude and direction of the force the electron initially experiences as it enters the magnetic field?



8 An alpha particle with a charge of $+3.2 \times 10^{-19}$ C is moving eastwards into a magnetic field acting into the page, as shown below. The force it experiences is *F*. If the velocity, *v*, of the particle is doubled, what will be the magnitude and direction of the magnetic force it would experience in terms of *F*?



9 The diagram below depicts a cross-sectional view of a long, straight, current-carrying conductor, located between the poles of a permanent magnet. The magnetic field, *B*, of the magnet, and the current, *I*, are perpendicular. Calculate the magnitude and direction of the magnetic force on a 5.0 cm section of the conductor when the current is 2.0 A into the page and *B* equals 2.0×10^{-3} T.



10 An east–west power line of length 80 m is suspended between two towers. Assume that the strength of the magnetic field of the Earth in this region equals 4.5×10^{-5} T.

a Calculate the magnitude and direction of the magnetic force on this power line at the moment it carries a current of 50 A from east to west.

b Over time, the ground underneath the eastern tower subsides, so that the power line is lower at that tower. Assuming that all other factors are the same, is the magnitude of the magnetic force on the power line greater than before, less than before or the same as before?

2.5 Comparing fields—a summary

Many of the forces affecting us and the world around us can be described as contact forces. There is direct contact as you open a door, kick a ball or rest on a couch. By contrast, the forces of gravity, magnetism and electricity act over a distance without necessarily having any physical contact (see Figure 2.5.1). This was a difficult idea for scientists to come to terms with. Newton still had some misgivings even when publishing his ideas of universal gravitation. The concept of fields, used to explain how and why forces can act over a distance, is thus a very powerful tool and one that has allowed us to better explain the fundamental forces of gravity and electromagnetism.

FIGURE 2.5.1 The magnet has an effect on the paper clips even though they are not in contact. This is because the paper clips are within the magnetic field produced by the magnet.

In this section key concepts and ideas on gravitational, electric and magnetic fields will be summarised, compared and contrasted to give an overview of the theoretical ideas covered so far.

DIPOLES AND MONOPOLES

Gravitational fields consist essentially of **monopoles**. All objects with mass produce a gravitational field that can be considered as being toward the centre of the mass. There is a concept of a gravitational dipole but it is a measure of how the mass of a single object is distributed away from a particular centre in a particular direction, usually selected as the centre of mass.

Magnetic fields exist in a practical sense solely as dipoles; that is, they have opposite north and south poles. While a magnetic field is defined as having a direction that a north magnetic pole would move (i.e. toward a south pole), this is a theoretical single pole.

Electric fields have both monopoles and dipoles. Single positive and negative charges represent monopoles. Two equal point charges of opposite sign separated by a distance, r, constitute a dipole. These exist often in physics and in areas such as molecular biology.

EXTENSION

Quadrupoles

Gravitational fields are said to also have a 'quadrupole'. A quadrupole is a representation of how the mass of an object is stretched out along a particular rotational axis. A sphere would thus have a zero quadrupole as the mass is evenly distributed around all axes. A long rod would have a quadrupole along its length. A flat plate would also have a quadrupole, but with the opposite sign of that of a rod since the axis would be pointing out either side of its flat sides rather than along its axis as for a rod. In general, quadrupoles can exist along x, y and z axes, each axis being at right angles to the others.

PHYSICSFILE

Gravitational repulsive forces

A leading theory in the explanation of the expansion of the universe is the concept of dark energy. While little is understood about dark energy at this time, it may be a source of a repulsive force of gravity possibly originating from the interaction of matter and antimatter.

DIRECTION AND SHAPE OF FIELDS

Simple fields associated with a single monopole, whether that be gravitational, electric or magnetic (although the magnetic one would be purely theoretical), look very similar since they are a representation of the spread of the field over the area being affected around a single point. Fields are vector quantities having both direction and size. Field lines are used to visualise the extent, shape and strength of the field, with arrows on the field lines used to show the direction of the field.

A uniform field would be indicated by lines that remain evenly spaced throughout the region of the field. The electric field in the region between two charged plates would be uniform. Around a point charge, mass or pole, while the field lines would be evenly spaced, the field would not be uniform since the strength of the field decreases with the distance from the charge. This is called a radial field.

In a static (unchanging) field, the strength of the field doesn't change with time. This is true of most gravitational and magnetic fields where the mass of the object or the strength of the magnet is unchanging. Many electric fields are changing fields. Charges are moving or the amount of charge is changing regularly with time. Of course there can be static electric fields with a fixed charge just as there can be changing gravitational and magnetic fields. The magnetic field associated with a changing electric current is one example of a changing magnetic field.

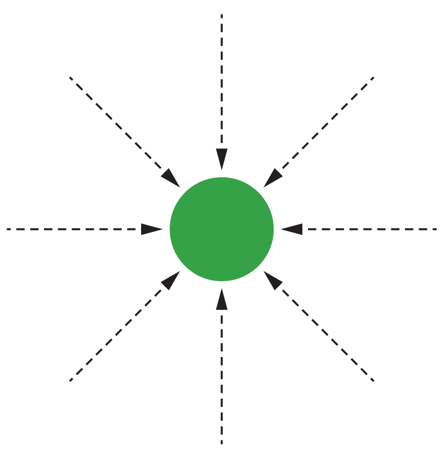


FIGURE 2.5.2 The arrows in this gravitational field diagram around a planet indicate that objects will be attracted towards the centre of mass of the planet. The spacing of the lines shows that force is stronger as you approach the planet. Very similar field diagrams apply to single electric point charges and magnetic poles.

A gravitational field is directed toward the point representing the centre of mass of the object and is always attractive (see Figure 2.5.2). In the case of both electric and magnetic fields, the field may be either attractive or repulsive so a particular direction is defined as the positive direction. In the case of electric fields this is the direction of the force on a positive test charge (i.e. positive to negative) and for magnetic fields, the direction of the force on a theoretical single north pole (i.e. north to south).

One other key difference between each of these fields is that theoretically a gravitational field around any mass extends an infinite distance from it. While the shape of the field will be influenced by the field of other masses, there is no way of stopping the field. The extent of both electric and magnetic fields, while theoretically extending to infinity, can be constrained by external electric and magnetic influences.

The shape of electrical and magnetic fields around objects can be influenced by the shape of the object. An example is shown in Figure 2.5.3. The shape of the field around multiple masses, charges or poles becomes increasingly complex. However, the direction of a field at any point is always the resultant field vector determined by adding the individual field vectors due to each mass, charge or magnetic pole within the affected region. Note, again, that gravitational fields are known only to be attractive.

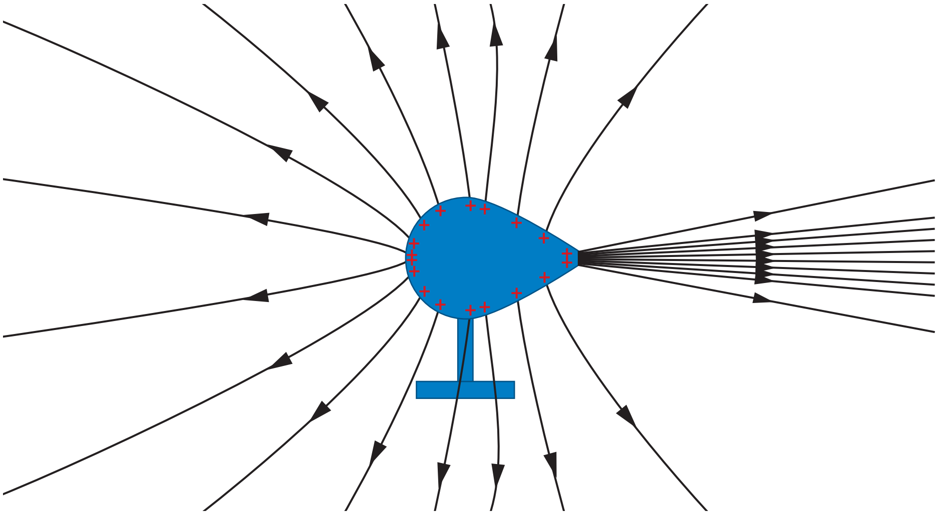


FIGURE 2.5.3 Field lines around a pear-shaped conductor. The uneven nature of their distribution is due to the contributions of each individual charge on the surface of the conductor and the greater density at more curved regions.

Whether a charge is positive or negative, or a magnetic pole is north or south, needs also to be considered when determining the resultant field around multiple charges or magnetic poles, as shown in Figure 2.5.4.

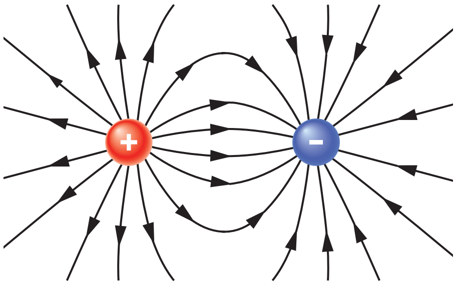


FIGURE 2.5.4 The electric field resulting when unlike charges are brought together. At any point the density and direction of the field lines represent the resultant field vector at that point.

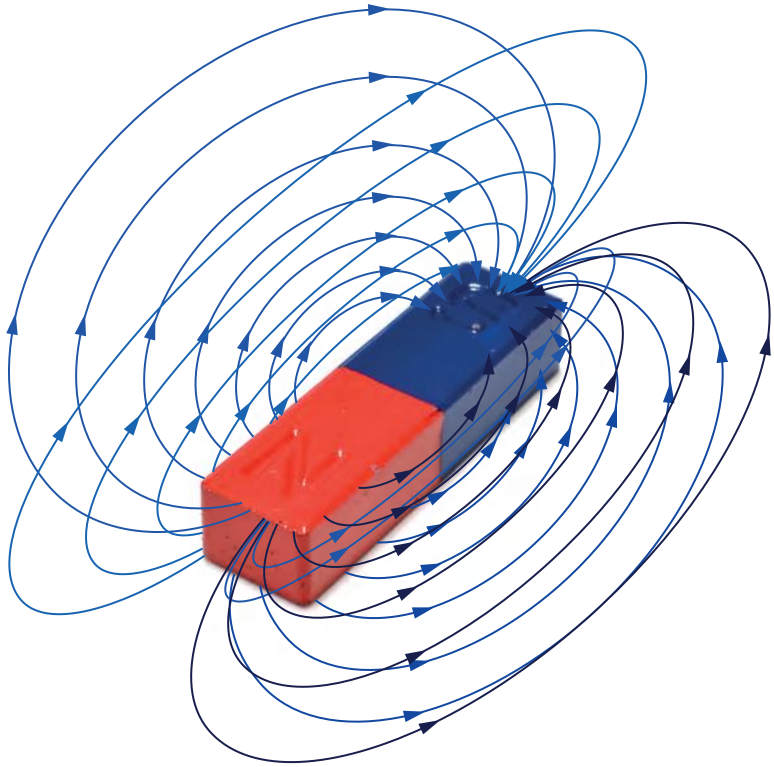


FIGURE 2.5.5 The magnetic field around a bar magnet. While fields are generally shown only in two dimensions, fields exist and affect the area around poles, charges and masses in all three dimensions.

PHYSICSFILE

Field strength around a dipole

While the theories of many particle physicists predict magnetic monopoles, in practical terms magnetic poles exist only as dipoles. The relevant relationships for dipoles are somewhat different to those for radial fields surrounding monopoles. An inverse square law, for example, does not apply. It can be shown that the field strength at a distance from a dipole will decrease with the cube of the distance. That is, field strength $\propto \frac{1}{r^3}$.

COMPARING GRAVITATIONAL AND ELECTRIC FIELDS

Gravity is an incredible force. Permeating the universe, it brings gas clouds together to form planets, stars and galaxies. It causes stars to collapse to black holes, generating gravitational fields strong enough that even light can't escape. And yet the gravitational force of attraction between two electrons is less than 8×10^{-37} N, which is the same as the electrostatic repulsion between the same two electrons.

The relationships developed for gravitational and electric fields over the last two chapters reveal the parallels and differences between related field concepts for gravitational masses and point charges, both of which are essentially monopoles. They are summarised in Table 2.5.1.

Quantity or description	Gravitational fields	Electrical fields
how field strength varies with distance, *r*, from a monopole	$g = G\frac{M}{r^2}$	$E = k\frac{Q}{r^2}$
force between monopoles	$F_g = G\frac{m_1 m_2}{r^2}$	$F = k\frac{q_1 q_2}{r^2}$
potential energy changes in a uniform field	$E_g = mg\Delta h$	$W = qV$
force due to a uniform field	$F_g = mg$	$F = qE$

TABLE 2.5.1 Comparison of gravitational and electric fields.

2.5 Review

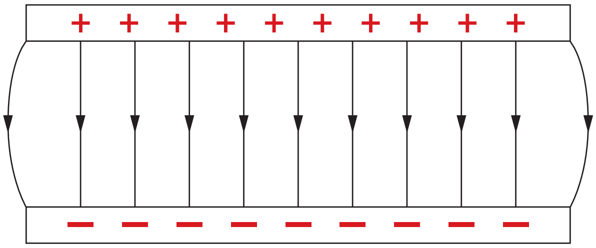
SUMMARY

- Gravitational, electric and magnetic fields are similar, but display significant differences associated with the differences in the fundamental nature of the fields.
- The direction of a field at any point is always the resultant field vector determined by adding the individual field vectors due to each mass, charge or magnetic pole within the affected region.
- A uniform field would be indicated by lines that remain evenly spaced throughout the region of the field.
- In a static (unchanging) field, the strength of the field doesn't change with time.
- The field around a monopole is radial, static but not uniform. It varies with the distance from the point source.

Quantity or description	Gravitational fields	Electrical fields	Magnetic fields
type of poles	monopoles	monopoles/dipoles	dipoles
type of force	attractive	attractive/repulsive	attractive/repulsive
extent of the field	extends to an infinite distance	can be constrained to a fixed distance	can be constrained to a fixed distance
effect of distance on field strength in a radial field	$g = G\frac{M}{r^2}$	$E = k\frac{Q}{r^2}$	
force between monopoles	$F_g = G\frac{m_1 m_2}{r^2}$	$F = k\frac{q_1 q_2}{r^2}$	
potential energy changes in a uniform field	$E_g = mg\Delta h$	$W = qV$	
force due to a uniform field	$F_g = mg$	$F = qE$	

KEY QUESTIONS

1 The diagram below shows the electric field between two electrically charged plates of opposite sign.



Choose the correct response that explains why the electric field lines shown are bulging outwards at the ends of the plates. Hint: Consider the field between the plates and how this compares with the field outside the plates. How would this affect the shape of the field at the ends of the plate?

A The plates are being drawn together by gravity, squeezing the electric field outwards at the ends.
B At either end a magnetic field is created, interacting with the electric field.
C At either end the horizontal component of the resultant force is outwards; between the plates it is zero.
D At either end air pressure will cause the field lines to bend outwards.

2 Determine which of the following statements is incorrect.

A Gravitational fields are known only to be attractive.
B In a static field, the strength of the field changes with time.
C Gravitational fields consist essentially of monopoles.
D Fields are vector quantities having both direction and size.

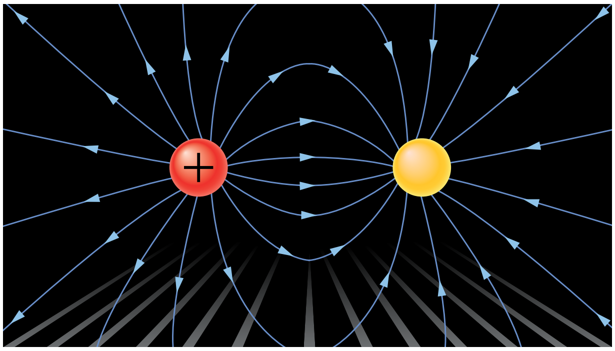
3 The gravitational force of attraction between two electrons is said to be less than 8×10^{-37} N.
At what minimum distance does this hold true for the two electrons?
(Use $G = 6.67 \times 10^{-11}$ m^3 kg^{-1} s^{-2}, and the mass of each electron is 9.1×10^{-31} kg.)

4 Describe the nature of the poles (monopoles, dipoles or both) for each of the following fields:

a gravitational
b electrical
c magnetic.

5 Complete the following statement about the field around a monopole from the pairs of choices provided in bold:
The field around a monopole is **linear/radial**, **static/dynamic** and **uniform/non-uniform**.

6 The diagram below shows the field between two point charges.
The charge on the right is shown with no sign. What is the charge on the point charge on the right?



7 Which of the following statements explains why an inverse square law does not apply to the change in magnetic field strength with distance from the source.

A Magnetic fields are considerably stronger than other field types.
B Magnetic fields are uniform around each pole.
C Magnetic fields are only associated with monopoles.
D Magnetic fields are only associated with dipoles.

8 Complete the following statement about the direction of a field around a monopole from the choices provided in bold:
The direction of a field at any point is defined as the **maximum/resultant** field vector determined by adding the **total/individual** field vectors due to each mass, charge or magnetic pole within the field.

9 The electron of a hydrogen atom orbits the single proton at the centre at an average distance of 0.53×10^{-10} m. The charges on both the electron and the proton are 1.6×10^{-19} C. What is the electrical force of attraction between the two particles?
(Use $k = 9 \times 10^9$ N m^2 C^{-2}.)

10 The electron of a hydrogen atom orbits the single proton at the centre at an average distance of 0.53×10^{-10} m. The mass of the electron is 9.1×10^{-31} kg and that of the proton is 1.67×10^{-27} kg. What is the gravitational force of attraction between the two particles?
(Use $G = 6.67 \times 10^{-11}$ m^3 kg^{-1} s^{-2}.)

Chapter review

02

KEY TERMS

cathode ray tube
dipole
electric field
electric field strength
electrical potential
field
field lines
Lorentz force
magnetic
magnetic field
magnetic pole
mnemonic
monopole
point charge
pole
potential difference
right-hand rules
solar wind
solenoid
voltaic pile

1 Calculate the force applied to an oil drop carrying a charge of 3.00 mC in a uniform electric field of 7.50 N C^{-1}.

2 A test charge is placed at a point, P, 30 cm directly above a charge, *Q*, of $+30 \times 10^{-6}$ C. What is the magnitude and direction of the electric field at point P?

A 300 N C^{-1} downwards
B 300 N C^{-1} upwards
C 3×10^6 N C^{-1} downwards
D 3×10^6 N C^{-1} upwards

3 Explain the difference between electrical potential and potential difference.

4 Calculate the potential difference that exists between two points separated by 25.0 mm, parallel to the field lines, in an electric field of strength 1000 V m^{-1}.

5 Between two plates forming a uniform electric field, where will the electrical field strength be at a maximum?

A close to the positive plate
B close to the earthed plate
C at all points between the plates
D at the mid-point between the plates

6 Choose the correct terms from the ones in bold to complete the relationship between work done and potential difference.
When a positively charged particle moves across a potential difference from a positive plate towards an earthed plate, work is done by the **field**/**charged particle** on the **field**/**charged particle**.

7 Calculate the work done to move a positively charged particle of 2.5×10^{-18} C a distance of 3.0 mm towards a positive plate in a uniform electric field of 556 N C^{-1}.

8 A particular electron gun accelerates an electron across a potential difference of 15 kV, a distance of 12 cm between a pair of charged plates. What is the magnitude of the force acting on the electron? (Use $q_e = 1.6 \times 10^{-19}$ C.)

9 A charge of $+q$ is placed a distance r from another charge also of $+q$. A repulsive force of magnitude F is found to exist between them. Choose the correct options from the ones in bold to describe the changes, if any, that will occur to the force in the following situations.

a The distance between the charges is doubled to $2r$, so the force will **halve**/**double**/**quadruple**/**quarter** and **repel**/**attract**.
b The distance between the charges is halved to $0.5r$, so the force will **halve**/**double**/**quadruple**/**quarter** and **repel**/**attract**.
c The distance between the charges is doubled and one of the charges is changed to $-2q$, so the force will **halve**/**double**/**quadruple**/**quarter** and **repel**/**attract**.

10 A gold(III) ion is accelerated by the electric field created between two parallel plates separated by 0.020 m. The ion carries a charge of $+3e$ and has a mass of 3.27×10^{-25} kg. A potential difference of 1000 V is applied across the plates. The work done to move the ion from one plate to the other results in an increase in the kinetic energy of the gold(III) ion. If the ion starts from rest, calculate its final velocity. (Use $q_e = -1.602 \times 10^{-19}$ C.)

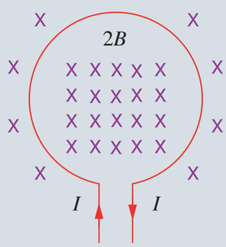
11 Calculate the magnitude of the force that would exist between two point charges of 5.00 mC and 4.00 nC separated by 2.00 m. (Use $k = 9 \times 10^9$ N m^2 C^{-2}.)

12 A point charge of 2.25 mC is positioned on top of an insulated rod on a table. At what distance above the point charge should a sphere of mass 3.00 kg containing a charge of 3.05 mC be located, so that it is suspended in the air?
(Use $k = 9 \times 10^9$ N m^2 C^{-2}.)

13 A charged plastic ball of mass 5.00 g is placed in a uniform electric field pointing vertically upwards with a strength of 300.0 N C^{-1}. Calculate the magnitude and sign of the charge required on the ball in order to create a force upwards that exactly equals the weight force of the ball.

The following information relates to questions 14–16.

The diagram below shows a loop carrying a current I that produces a magnetic field of magnitude B in the centre of the loop. It is in a region where there is already a steady field of magnitude B (the same magnitude as that due to I) directed into the page. The resultant magnetic field has a magnitude of $2B$.



14 What would the magnitude and direction of the resultant field be at the centre of the loop if the current in the loop is switched off?

15 What would the magnitude and direction of the resultant field at the centre of the loop be if the current in the loop were doubled?

16 What would the magnitude and direction of the resultant field at the centre of the loop be if the current in the loop were reversed but maintained the same magnitude?

17 Complete the following sentence by selecting the best option.

The magnitude of the magnetic force on a conductor aligned so that the current is running parallel to a magnetic field is:

A dependent on the size of the current

B dependent on the size of the magnetic field

C dependent on the length of the conductor

D zero

E a maximum

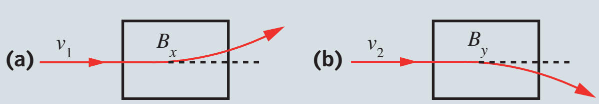
18 The right-hand rule is used to determine the force on a current-carrying conductor perpendicular to a magnetic field. Identify what part of the hand corresponds to the following physical quantities:

a magnetic force

b magnetic field

c current in the conductor.

19 The following diagrams (a) and (b) show two different electron beams being bent as they pass through two different regions of a uniform magnetic field of equal magnitudes B_x and B_y. The initial velocities of the electrons in the respective beams are v_1 and v_2. Complete the following sentence by choosing the correct term from those in **bold**.



For the electron beams to behave as shown in (a), v_1 is **equal to/less than** v_2 and the region of the magnetic field, B_y, must be acting **out of/into** the page.

20 How much current, I, must be flowing in a wire 3.2 m long if the maximum force on it is 0.800 N and it is placed in a uniform magnetic field of 0.0900 T?

21 Calculate the magnitude and direction of the magnetic force on conductors with the following sets of data:

a B = 1.0 mT left, l = 5.0 mm, I = 1.0 mA up

b B = 0.10 T left, l = 1.0 cm, I = 2.0 A up

22 Calculate the force exerted on an electron ($q = 1.6 \times 10^{-19}$ C) travelling at a speed of 7.0×10^{6} m s^{-1} at right angles to a uniform magnetic field of strength 8.6×10^{-3} T.

23 A horseshoe magnet is held vertically with the north pole of the magnet on the left and the south pole of the magnet on the right.

What is the direction of the magnetic force acting on the wire?

24 Power lines carry an electric current in the Earth's magnetic field. Which would experience the greater magnetic force: a north–south power line or an east–west power line? Explain your answer.

25 Which of the following types of fields would you NOT expect to be associated with radial fields?

A gravitational

B electrical

C magnetic

D all of the above

26 Two electrons approach each other at a distance of 5.4×10^{-12} m. The charge on both the electrons is -1.6×10^{-19} C. What is the electrical force of repulsion between the two electrons? (Use $k = 9 \times 10^{9}$ N m^2 C^{-2}.)

27 Two electrons approach each other at a distance of 5.4×10^{-12} m. The mass of each electron is 9.1×10^{-31} kg. What is the gravitational force of attraction between the two electrons? (Use $G = 6.67 \times 10^{-11}$ m^3 kg^{-1} s^{-2}.)

CHAPTER 03

Applications of fields

As explained in the previous chapters, gravitational, magnetic and electric fields affect things that are some distance away. There does not need to be direct contact for fields to exert a force.

This chapter looks at the application of these gravitational, magnetic and electric fields. You will use your understanding of fields to explain how DC motors operate, to understand satellite motion and to predict how charged particles will behave in electric fields.

Key knowledge

At the end of this chapter, you will have studied the application of fields and will be able to:

- analyse the use of a magnetic field to change the path of a charged particle, including the radius of the path followed by a low-velocity electron in a magnetic field:
  $qvB = \frac{mv^2}{r}$
- apply the concepts of force due to gravity, F_g, and normal reaction force, F_N, including satellites in orbit where the orbits are assumed to be uniform and circular
- model satellite motion (artificial, Moon, planet) as uniform circular orbital motion:
  $a = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2}$
- investigate and analyse theoretically and practically the operation of simple DC motors consisting of one coil, containing a number of loops of wire, which is free to rotate about an axis in a uniform magnetic field and including the use of a split ring commutator
- model the acceleration of particles in a particle accelerator (limited to linear acceleration by a uniform electric field and direction change by a uniform magnetic field).

3.1 Satellite motion

When Isaac Newton developed his law of universal gravitation, as discussed in Chapter 1 'Gravity', he was building on work previously done by Nicolaus Copernicus, Johannes Kepler and Galileo Galilei. Copernicus had proposed a sun-centred (heliocentric) solar system. Galileo had developed laws relating to motion near the Earth's surface and Kepler had devised rules concerned with the motion of the planets. Kepler published his laws on the motion of planets 80 years before Newton published his law of universal gravitation.

In this section, you will look at how Newton synthesised the work of Galileo and Kepler and proposed that the force that was causing an apple to fall to the Earth was the same force that was keeping the Moon in its orbit. Newton was the first to propose that satellites could be placed in orbit around Earth, almost 300 years before it was technically possible to do this. Now, thousands of artificial satellites are in orbit around Earth and are an essential part of modern life (see Figure 3.1.1).

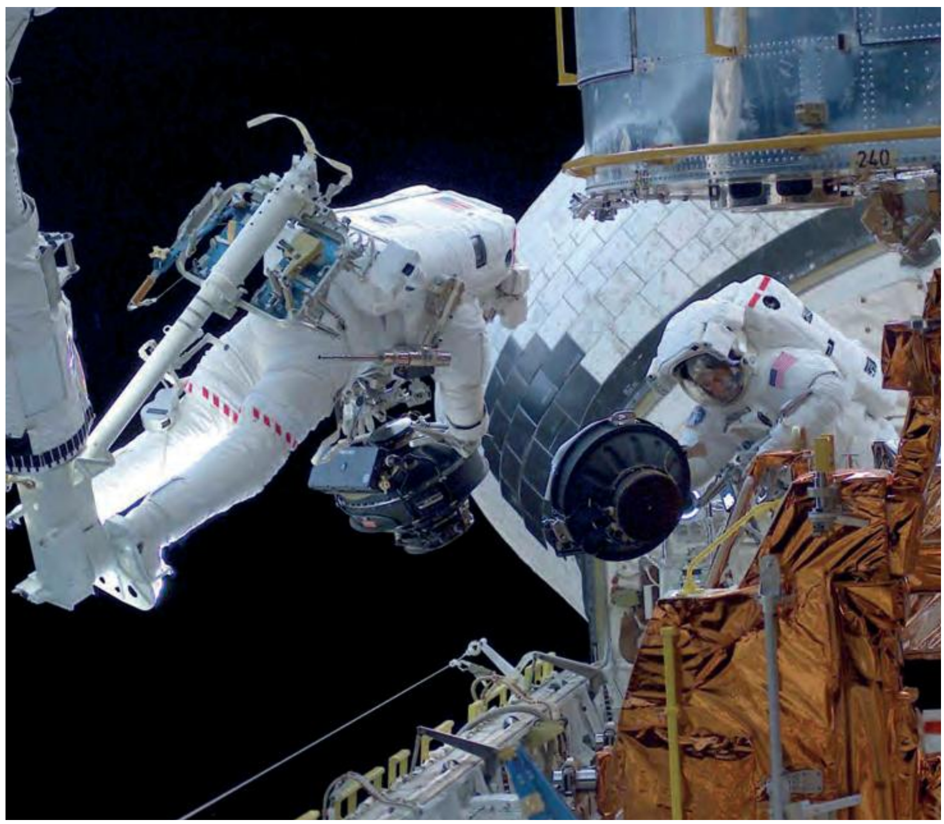


FIGURE 3.1.1 Astronauts on a repair mission to the Hubble Space Telescope (HST) in 1994. The satellite initially malfunctioned, but the repair was successful and the HST is still going strongly.

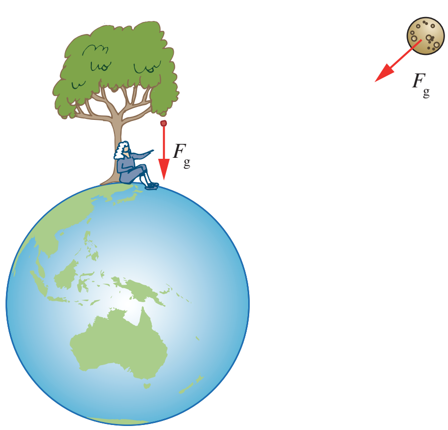


FIGURE 3.1.2 Newton realised that the gravitational attraction of the Earth (F_g) was determining the motions of both the Moon and the apple.

NEWTON'S THOUGHT EXPERIMENT

A **satellite** is an object in a *stable orbit* around another object. Isaac Newton developed the notion of satellite motion while working on his theory of gravitation. He was comparing the motion of the Moon with the motion of a falling apple and realised that it was the gravitational force of attraction towards the Earth that determined the motion of both objects (see Figure 3.1.2). He reasoned that if this force of gravity was not acting on the Moon, the Moon would move at constant speed in a straight line at a tangent to its orbit.

Newton proposed that the Moon, like the apple, was also falling. It was continuously falling to the Earth without actually getting any closer to the Earth. He devised a thought experiment in which he compared the motion of the Moon with the motion of a cannonball fired horizontally from the top of a high mountain.

His thought experiment is illustrated in Figure 3.1.3. In this thought experiment, if the cannonball was fired at a low speed, it would not travel a great distance before gravity pulled it to the ground (see the shortest dashed line in Figure 3.1.3(b)). If it was fired with a greater velocity, it would follow a less curved path and land a greater distance from the mountain (see the next two dashed lines in Figure 3.1.3(b)). Newton reasoned that, if air resistance was ignored and if the cannonball was fired fast enough, it could travel around the Earth and reach the place from where it had been launched (shown by the solid circular line in Figure 3.1.3(b)). At this speed, it would continue to circle the Earth indefinitely even though the cannonball has no propulsion system.

In reality, satellites could not orbit the Earth at low altitudes, because of air resistance. Nevertheless, Newton had proposed the notion of an artificial satellite hundreds of years before one was actually launched. Any object placed at the right altitude with enough speed would simply continue in its orbit.

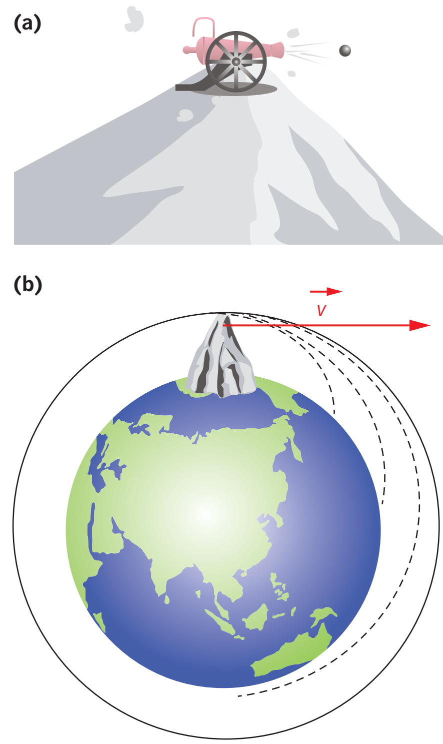


FIGURE 3.1.3 These diagrams show how a projectile that was fired fast enough from a very high mountain (a) would fall all the way around the Earth and become an Earth satellite (b).

MASS AND WEIGHT

In Unit 2 Physics, the concept of weight force was introduced. You will recall that the force of weight is another name for the gravitational force on an object when the object is near the surface of a planet or other large body, e.g. the Earth or the Moon.

Weight force (F_g or W) is equal to the mass (m) of the object multiplied by the acceleration due to gravity (g) at the place you are measuring the weight. Therefore:

$$F_g = mg$$

where F_g is the force due to gravity or weight (N)

m is mass of the object (kg)

g is the gravitational field strength (N kg^{-1})

On the Earth, the value of g is taken as 9.80665 m s^{-2} or 9.80665 N kg^{-1}. This is often rounded to 9.8 m s^{-2}, or in some cases is approximated as 10 m s^{-2} for ease of calculations.

The weight force or weight of an object is measured in newtons and, because it is a vector quantity, requires a direction. The weight of an object is always downwards, towards the centre of the planet or moon, etc. (e.g. Earth).

NORMAL FORCE

If you exert a force against a wall, Newton's third law says that the wall will exert an equal but opposite force on you. If you push with a greater force, the wall will also exert a greater force. This is shown in Figure 3.1.4.

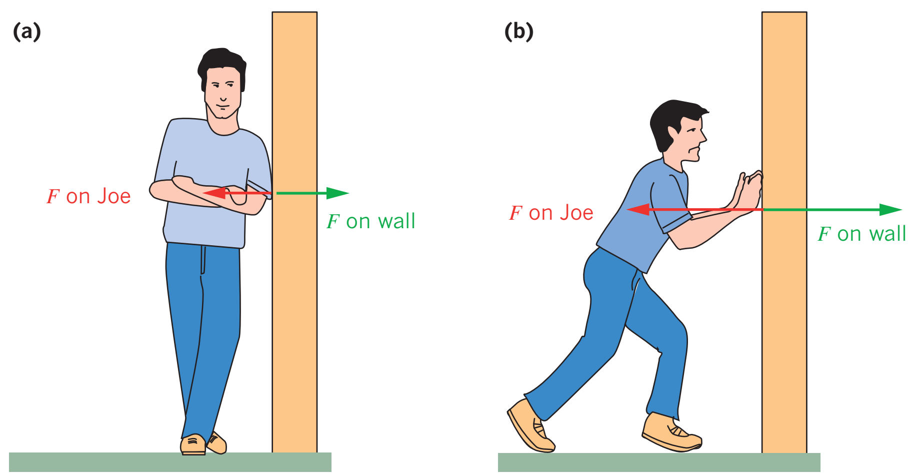


FIGURE 3.1.4 (a) If Joe exerts a small force on the wall, the wall exerts a small force on Joe. (b) When Joe pushes hard against the wall, the wall pushes back just as hard! In both (a) and (b), the red and green arrows are equal in size but opposite in direction. That is, $F_{\text{on Joe by wall}} = -F_{\text{on wall by Joe}}$.

The force from the wall acts at right angles to the surface, i.e. it is *normal* to the surface and is thus called a normal force. Like every force, a normal force is one half of an action/reaction pair, so it is often called a **normal reaction force**. The normal force is represented by F_N or N.

For an object at rest on the ground, the normal force will be equal in size to the weight force of the object.

During many interactions and collisions, the size of the normal force changes. For example, when a ball bounces, the forces that act on it during its contact with the floor are its weight, F_g, and the normal force, F_N, from the floor. The series of diagrams shown in Figure 3.1.5 depict the changes in the magnitude of the normal force throughout the bounce of a ball.

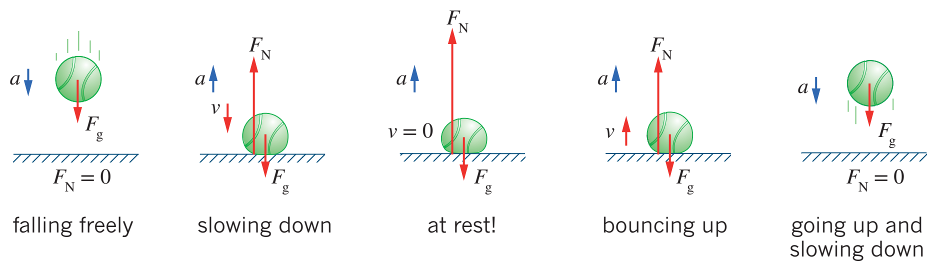


FIGURE 3.1.5 The forces acting on a bouncing ball before, during and after striking the floor.

When contact has just been made, the ball is compressed only slightly, indicating that the force from the floor is minimal. This force then becomes larger and larger, causing the ball to become more and more deformed. At the point of maximum compression, the normal force is at its maximum value and the bouncing ball is momentarily stationary.

The forces acting on a ball as it bounces (its weight, F_g, and the normal force, F_N) are not an action/reaction pair. Both *act on the same body*, whereas Newton's third law describes forces that bodies exert on *each other*. A pair of action/reaction forces that act during the bounce are the upwards force, F_N, that the floor exerts on the ball and the downwards force that the ball exerts on the floor (not shown in Figure 3.1.5). This downwards force is equal in magnitude to the normal force, so it too varies during the bounce.

Apparent weight

Your **apparent weight** is the same size as the normal reaction force that acts upwards on your feet from a surface. It results from your weight force pulling you downwards onto the floor. The reason why the upwards reaction force is called your 'apparent' weight is because you do not feel the force you apply to the floor, you will only experience with your senses the forces that are applied on *you*. What you feel is the normal force acting up on you from the floor. Normally, when you stand on a surface that is either stationary or in constant vertical motion, your apparent weight is constant and equal to your weight force (see Figure 3.1.6).

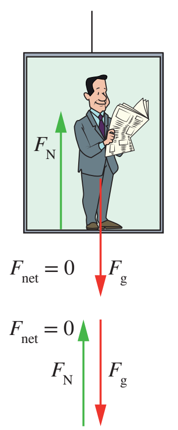


FIGURE 3.1.6 In this case, the forces that act on the person, F_N and F_g, are equal in size. The person will 'feel' his normal apparent weight.

The apparent weight that you experience changes when the surface you are standing on is accelerating upwards or downwards. If the floor is accelerating downwards at a rate less than 9.80 m s^{-2}, your feet will be pressing less firmly on the surface than when the floor was not accelerating. Therefore, the normal force is also less and so your apparent weight appears to be less. That is, you would feel lighter than usual (see Figure 3.1.7).

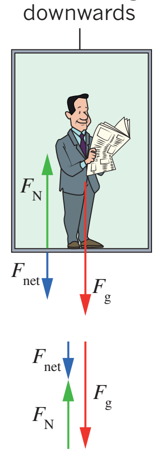


FIGURE 3.1.7 In this case, the forces that act on the person in the lift cause him to feel lighter than his normal apparent weight. When accelerating downwards, $F_N < F_g$.

The opposite happens when the floor is accelerating upwards. In this case, the floor is pushing up against your feet with a greater force than the normal reaction force due to your weight alone. The upwards push of the floor must provide the force to accelerate you upwards. This accelerating force adds to the normal force to make it appear that your apparent weight is greater than it would be if you weren't accelerating. That is, you would feel heavier than usual (see Figure 3.1.8).

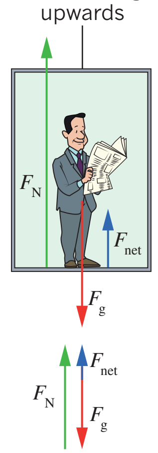


FIGURE 3.1.8 In this case, the forces that act on the person in the lift cause him to feel heavier than his normal apparent weight. When accelerating upwards, $F_N > F_g$.

The normal reaction force (felt as apparent weight) and the force due to gravity (weight force) add as vectors to give the net force that causes the acceleration:

$F_{net} = F_N + F_g$

where F_N is the apparent weight force that acts upwards on your feet

F_g is the weight force due to gravity (which never changes)

F_{net} is the net force causing the acceleration

Worked example 3.1.1

CALCULATING APPARENT WEIGHT

A 79.0 kg student rides a lift up to the top floor of an office block. During the journey, the lift accelerates upwards at 1.26 m s^{-2} before travelling at a constant velocity of 3.78 m s^{-1} and then finally decelerating at 1.89 m s^{-2}.

a Calculate the apparent weight of the student in the first part of the journey while accelerating upwards at 1.26 m s^{-2}.

Thinking	Working
Ensure that the variables are in their standard units.	m = 79.0 kg a = 1.26 m s^{-2} up g = 9.80 m s^{-2} down
Apply the sign and direction convention for motion in one dimension. Up is positive and down is negative.	m = 79.0 kg a = +1.26 m s^{-2} g = −9.80 m s^{-2}
Apply the equation for apparent weight (the normal force).	$F_{net} = F_N + F_g$ $F_N = F_{net} - F_g$ $= ma - mg$ $= (79.0 \times 1.26) - (79.0 \times -9.80)$ $= 99.54 + 774.2$ = 874 N

b Calculate the apparent weight of the student in the second part of the journey while travelling at a constant speed of 3.78 m s^{-1}.

Thinking	Working
Ensure that the variables are in their standard units.	m = 79.0 kg a = 0 m s^{-2} g = 9.80 m s^{-2} down
Apply the sign and direction convention for motion in one dimension. Up is positive and down is negative.	m = 79.0 kg a = 0 m s^{-2} g = −9.80 m s^{-2}
Apply the equation for apparent weight (the normal force).	$F_{net} = F_N + F_g$ $F_N = F_{net} - F_g$ $= ma - mg$ $= (79.0 \times 0) - (79.0 \times -9.80)$ $= 0 + 774.2$ = 774 N

c Calculate the apparent weight of the student in the last part of the journey while travelling upwards and decelerating at 1.89 m s^{-2}.	
Thinking	**Working**
Ensure that the variables are in their standard units. Also consider that deceleration is a negative acceleration.	m = 79.0 kg a = −1.89 m s^{-2} up g = 9.80 m s^{-2} down
Apply the sign and direction convention for motion in one dimension. Up is positive and down is negative.	m = 79.0 kg a = −1.89 m s^{-2} g = −9.80 m s^{-2}
Apply the equation for apparent weight or the normal force.	$F_{net} = F_N + F_g$ $F_N = F_{net} - F_g$ $= ma - mg$ $= (79.0 \times -1.89) - (79.0 \times -9.80)$ $= -149.3 + 774.2$ = 625 N

Worked example: Try yourself 3.1.1

CALCULATING APPARENT WEIGHT

A 79.0 kg student rides a lift down from the top floor of an office block to the ground. During the journey the lift accelerates downwards at 2.35 m s^{-2}, before travelling at a constant velocity of 4.08 m s^{-1} and then finally decelerating at 4.70 m s^{-2}.

a Calculate the apparent weight of the student in the first part of the journey while accelerating downwards at 2.35 m s^{-2}.

b Calculate the apparent weight of the student in the second part of the journey while travelling at a constant speed of 4.08 m s^{-1}.

c Calculate the apparent weight of the student in the last part of the journey while travelling downwards and decelerating at 4.70 m s^{-2}.

From these Worked examples, you can see that:

- when accelerating upwards the student will feel heavier than normal ($F_N > mg$) (Note: this is the same as decelerating while travelling downwards)
- when accelerating downwards, the student will feel lighter than normal ($F_N < mg$) (Note: this is the same as decelerating while travelling upwards)
- when travelling upwards or downwards at a constant velocity, the student will feel their normal weight, just as they would if the lift was stationary ($F_N = mg$).

Apparent weightlessness

Defining apparent weight makes it possible to identify the situations in which you will experience **apparent weightlessness**. Your apparent weight is a contact reaction force that acts upwards on you from a surface because gravity is pulling you down on that surface. So if you are not standing on a surface, then you will experience zero apparent weight or apparent weightlessness. This means that you will experience apparent weightlessness the moment you step off the top platform of a diving pool or as you skydive from a plane, although the rushing air will hardly let you experience the sensation of floating as you skydive.

Felix Baumgartner experienced apparent weightlessness as he fell from his balloon 39 kilometres above the Earth (see Figure 3.1.9). This vertical height is equivalent to the widest part of Port Phillip Bay.

FIGURE 3.1.9 Felix Baumgartner experienced apparent weightlessness on his return to Earth from 39 000 m.

Astronauts also experience apparent weightlessness in the International Space Station, which orbits about 370 kilometres above the surface of the Earth (about the horizontal distance from Melbourne to the town of Orbost).

Whenever you are in **free fall**, you experience apparent weightlessness. It follows then that whenever you experience apparent weightlessness, you must be in free fall. When astronauts experience apparent weightlessness, they are not floating in space as they orbit the Earth. They are actually in free fall. Astronauts and their spacecraft are both falling, but not directly towards the Earth like Baumgartner. The astronauts are actually moving horizontally, as shown in Figure 3.1.10. Baumgartner stayed approximately above the same place on the Earth from where he departed. Astronauts, on the other hand, are moving at a velocity relative to the Earth so they are moving across the sky at the same time as they are falling. The combined effect is that they fall in a curved path that exactly mirrors the curve of the Earth. So they fall, but continually miss the Earth as the surface of the Earth curves away from their path.

Importantly there is a significant difference between apparent weightlessness and true weightlessness. True weightlessness only occurs when the gravitational field strength is zero and hence $F_g = 0$. This only occurs in deep space, far enough away from any planets that their gravitational effect is zero. Apparent weightlessness, however, can occur when still under the influence of a gravitational field.

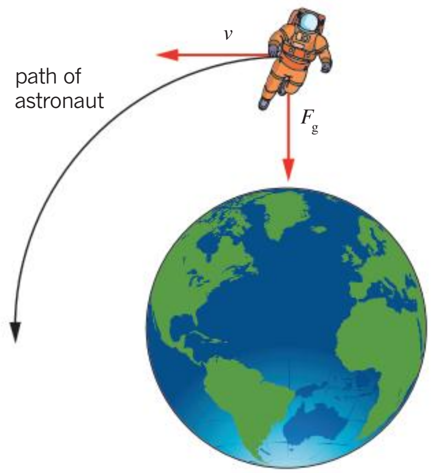


FIGURE 3.1.10 Astronauts are in free fall while orbiting the Earth.

EXTENSION

Falling at constant speed

Galileo was able to show more than 400 years ago that the mass of a body does not affect the rate at which it falls towards the ground. However, our common experience is that not all objects behave in this way. A light object, such as a feather or a balloon, does not accelerate at 9.80 m s^{-2} as it falls. It drifts slowly to the ground, far slower than other dropped objects. Parachutists and skydivers also eventually fall with a constant speed. However, they can change their falling speed by changing their body profile, as pictured in Figure 3.1.11. If they assume a tuck position, they will fall faster and if they spread out their arms and legs, they will fall slower. This enables them to form spectacular patterns as they fall.

FIGURE 3.1.11 Skydivers performing intricate manoeuvres in free fall.

Skydivers, base-jumpers and air-surfers are able to use the force of air resistance to their advantage. As a skydiver first steps out of their plane, the forces acting on them are drag (air resistance), F_{ar}, and weight due to gravity, F_g. Since their speed is low, the drag force is small as shown in Figure 3.1.12(a). There is a large net force (F_{net}) downwards, so they will experience a large downwards acceleration of just less than 9.80 m s^{-2}, causing them to speed up. This causes the drag force to increase because they are colliding harder with the air molecules. In fact, the drag force increases in proportion to the square of the speed, $F_{ar} \propto v^2$. This results in a smaller net force downwards as shown in Figure 3.1.12(b). Their downwards acceleration is therefore reduced. It is important to remember that they are still speeding up, but at a reduced rate.

As their speed continues to increase, so too does the magnitude of the drag force. Eventually, the drag force becomes as large as the weight force due to gravity, as shown in Figure 3.1.12(c). When this happens, the net force is zero and the skydiver will fall with a constant velocity. Since the velocity is now constant, the drag force will also remain constant and the motion of the skydiver will not change, as shown in Figure 3.1.12(d). This velocity is commonly known as the *terminal velocity*.

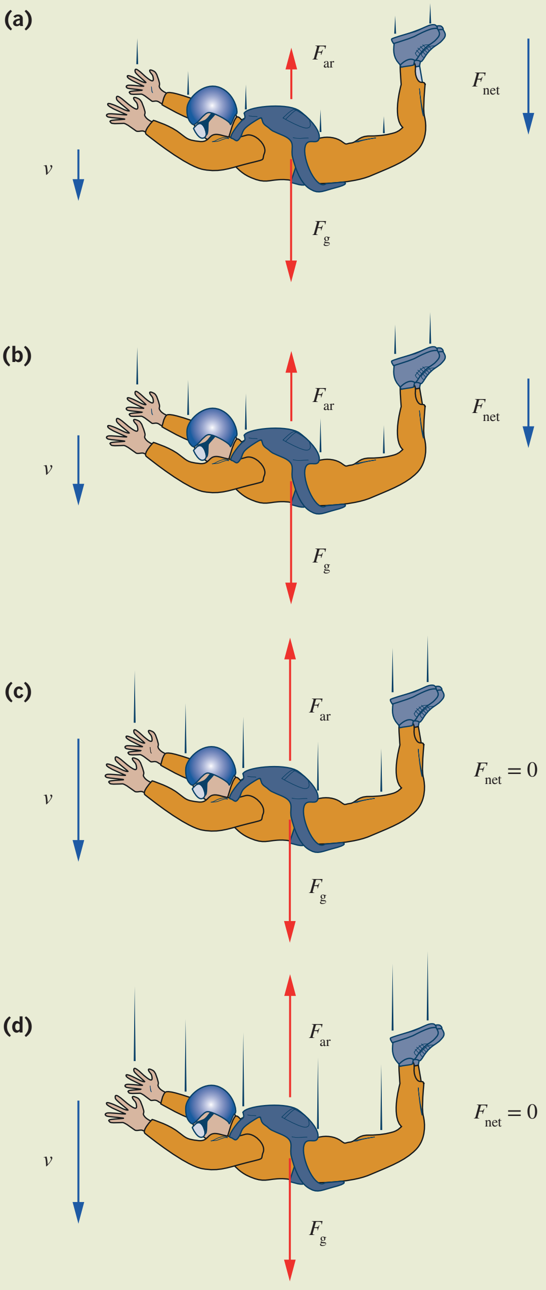


FIGURE 3.1.12 The forces involved in skydiving.

Natural satellites

Natural satellites have existed throughout the universe for billions of years. The planets and asteroids of the solar system are natural satellites of the Sun (see Figure 3.1.13).

The Earth has one natural satellite: the Moon. The largest planets—Jupiter and Saturn—have more than sixty natural satellites each in orbit around them. Most of the stars in the Milky Way galaxy have planets and more of these exoplanets are being discovered each year.

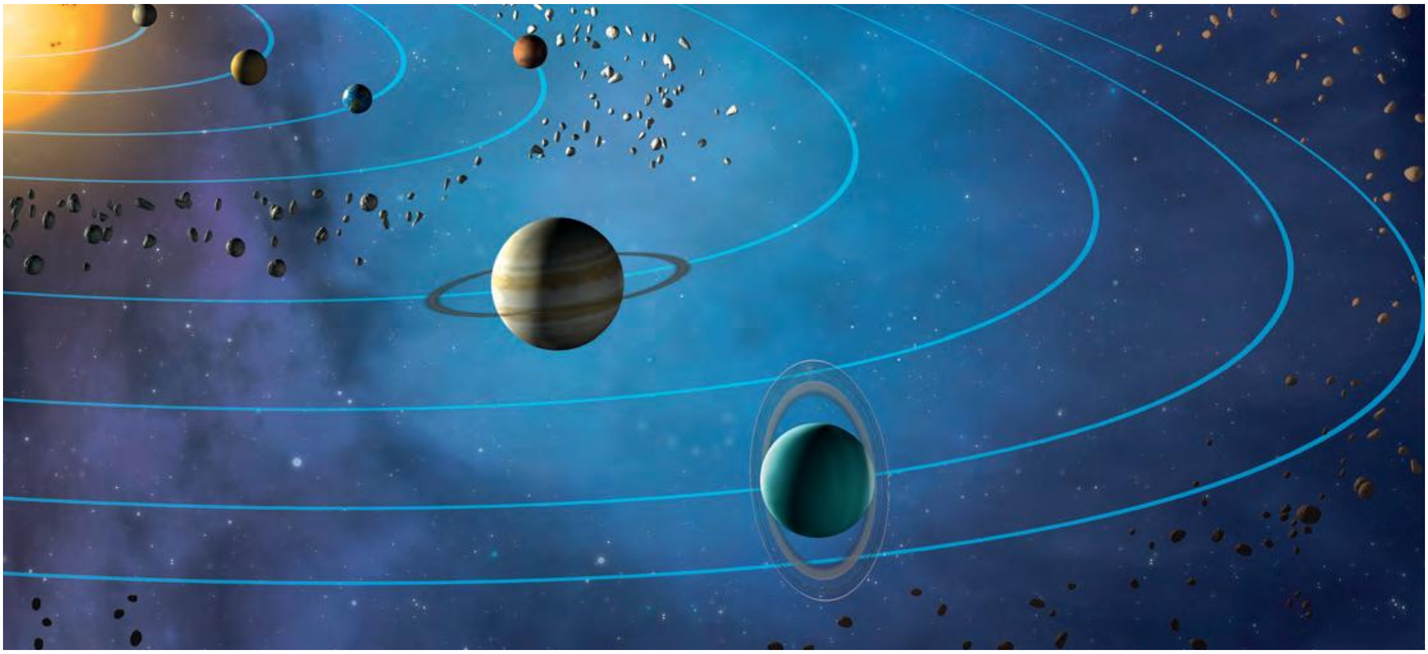


FIGURE 3.1.13 The planets are natural satellites of the Sun. The planets closer to the Sun have a shorter orbital period than the larger gas giants.

ARTIFICIAL SATELLITES

Since the Space Age began in 1957 with the launch of Sputnik, about 6000 **artificial satellites** have been launched into orbit around the Earth. Today there are around 4000 still in orbit, although only around 1200 of these are operational.

Satellites in orbit around the Earth are classified as low, medium or high orbit.

- Low orbit: 180 km to 2000 km altitude. Most satellites orbit in this range (an example is shown in Figure 3.1.14). These include the Hubble Space Telescope, which is used by astronomers to view objects right at the edge of the universe.
- Medium orbit: 2000 km to 36 000 km altitude. The most common satellites in this region are the Global Positional System (GPS) satellites used to run navigation systems.
- High orbit: 36 000 km altitude or greater. Australia uses the Optus satellites for communications, and deep-space weather pictures come from the Japanese MTSAT-1R satellite. The satellites that sit at an altitude of 36 000 km and orbit with a period of 24 hours are known as **geostationary satellites** (or geosynchronous satellites). Most communications satellites are geostationary.

Earth satellites can have different orbital paths depending on their function:

- equatorial orbits, where the satellite always travels above the equator
- polar or near-polar orbits, where the satellite travels over or close to the North and South Poles as it orbits
- inclined orbits, which lie between equatorial and polar orbits.

Satellites are used for a multitude of different purposes, with 60 per cent used for communications. Many low-orbit American NOAA satellites have an inclination of 99° and an orbit that allows them to pass over each part of the Earth at the same time each day. These satellites are also known as Sun-synchronous satellites.

FIGURE 3.1.14 A low-orbit satellite called the Soil Moisture and Ocean Salinity (SMOS) probe was launched in August 2014. Its role is to measure water movements and salinity levels on Earth as a way of monitoring climate change. It was launched from northern Russia by the European Space Agency (ESA).

PHYSICS IN ACTION

Three satellites

Geostationary Meteorological Satellite MTSAT-1R

The Japanese MTSAT-1R satellite was launched in February 2005, and orbits at 35 800 km directly over the equator. At its closest point to the Earth, known as the perigee, its altitude is 35 776 km. At its furthest point from the Earth, known as the apogee, it is at 35 798 km. MTSAT-1R orbits at a longitude of 140° E, so it is just to the north of Cape York and ideally located for use by Australia's weather forecasters. It has a period of 24 hours, so is in a geostationary orbit.

Signals from MTSAT-1R are transmitted every 2 hours and are received by a satellite dish on the roof of the head office at the Bureau of Meteorology in Perth. Infrared images show the temperature variations in the atmosphere and are invaluable in weather forecasting. MTSAT-1R is box-like and measures about 2.6 m along each side. It has a mass of 1250 kg and is powered by solar panels that, when deployed, take its overall length to over 30 m.

Hubble Space Telescope (HST)

This cooperative venture between NASA and the European Space Agency (ESA) was launched by the crew of the space shuttle Discovery on 25 April 1990. Hubble is a permanent unoccupied space-based observatory with a 2.4 m–diameter reflecting telescope, spectrographs and a faint-object camera. It orbits above the Earth's atmosphere, producing images of distant stars and galaxies far clearer than those from ground-based observatories (see Figure 3.1.15). The HST is in a low-Earth orbit inclined at 28° to the equator. Its expected life span was originally around 15 years, but service and repair missions have extended its life and it is still in use today.

National Oceanic and Atmospheric Administration Satellite (NOAA-19)

Many of the US-owned and operated NOAA satellites are located in low-altitude near-polar orbits. This means that they pass close to the poles of the Earth as they orbit. NOAA-19 was launched in February 2009 and orbits at an inclination of 99° to the equator. Its low altitude means that it captures high-resolution pictures of small bands of the Earth. The data is used in local weather forecasting as well as to provide enormous amounts of information for monitoring global warming and climate change.

Table 3.1.1 provides data for the three satellites discussed in this section.

FIGURE 3.1.15 In August 2014, astronomers used the Hubble Space Telescope to detect the blue companion star of a white dwarf in a distant galaxy. The white dwarf slowly siphoned fuel from its companion, eventually igniting a runaway nuclear reaction in the compact star, which produced a supernova blast.

Satellite	Orbit	Inclination	Perigee (km)	Apogee (km)	Period
MTSAT-1R	equatorial	0°	35 776	35 798	1 day
Hubble	inclined	28°	591	599	96.6 min
NOAA-19	near polar	99°	846	866	102 min

TABLE 3.1.1 A comparison of the three satellites discussed in this section.

PHYSICSFILE

SuitSat1

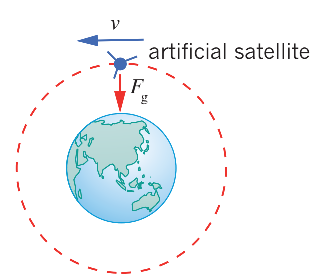
One of the more unusual satellites was launched from the International Space Station on 3 February, 2006. It was an obsolete Russian spacesuit into which the astronauts had placed a radio transmitter, batteries and some sensors. Its launch involved simply being pushed off by one of the astronauts while on a spacewalk. SuitSat1 was meant to transmit signals that would be picked up by ham radio operators on Earth for a few weeks, but transmissions ceased after just a few hours (see Figure 3.1.16). The spacesuit burned up in the atmosphere over Western Australia in September 2006.

SuitSat2 was launched in August 2011 and contained experiments created by school students. It re-entered Earth's atmosphere in January 2012 after 5 months in orbit.

FIGURE 3.1.16 This photograph does not show an astronaut drifting off to certain death in space. This is SuitSat1, one of the strangest satellites ever launched, at the start of its mission.

Artificial and natural satellites are not propelled by rockets or engines. They orbit in free fall and the only force acting on them is the gravitational attraction between themselves and the body about which they orbit. This means that the satellites have a **centripetal acceleration** that is equal to the gravitational field strength at their location (see Figure 3.1.17). Centripetal acceleration is covered in more detail in Chapter 5 'Newtonian theories of motion'.

Artificial satellites are often equipped with tanks of propellant that are squirted in the appropriate direction when the orbit of the satellite needs to be adjusted.



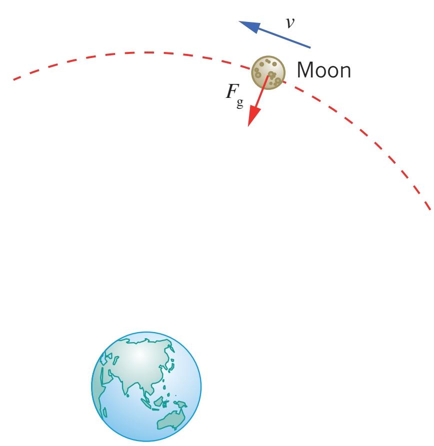


FIGURE 3.1.17 The only force acting on these artificial and natural satellites is the gravitational attraction of the Earth. Both orbit with a centripetal acceleration equal to the gravitational field strength at their locations.

PHYSICSFILE

See the International Space Station (ISS) and other satellites

It is easy to see low-orbit satellites if you are away from city lights. The best time to look is just after sunset. If you can, go outside and look for any slow moving objects passing across the star background.

There are also many websites that will allow you to track and predict the real-time paths of satellites. You can use the NASA 'Spot the Station' website to see when the ISS is passing over your part of the planet. The ISS is so bright that it is easy to see from most locations.

PHYSICSFILE

Space junk

Today there are around 1200 satellites that are still in operation. There are also around 2800 satellites that have reached the end of their operational life or have malfunctioned but are still in orbit.

In 2007, a Chinese satellite was deliberately destroyed by a missile, creating thousands of pieces of debris. In 2009, a collision between the defunct Russian Cosmos 2251 and operational US Iridium 33 created even more debris. This debris and the defunct satellites are classified as space junk (see Figure 3.1.18).

The presence of this fast-moving space junk puts the other satellites and the International Space Station at risk from collision. Currently around 22 000 pieces of space junk are being tracked and monitored. There have been a number of occasions where satellites have been moved to avoid collisions with space junk.

The UN has passed a resolution to remove defunct satellites from low-Earth orbits by placing them in much higher orbits, or bringing them back to Earth and allowing them to burn up in the atmosphere.

FIGURE 3.1.18 An exaggerated map showing the location of space debris and abandoned satellites in near-Earth orbits.

FIGURE 3.1.19 Johannes Kepler, who was the first to work out that the planets do not travel in circular paths, but rather in elliptical paths.

KEPLER'S LAWS

Kepler, a German astronomer (depicted in Figure 3.1.19), published his three laws regarding the motion of planets in 1609. This was about 80 years before Newton's law of universal gravitation was published. Kepler was analysing the motion of the planets in orbit around the Sun, but these laws can be used for any satellite in orbit around any central mass.

Kepler's laws are as follows:

1. The planets move in elliptical orbits with the Sun at one focus.
2. The line connecting a planet to the Sun sweeps out equal areas in equal intervals of time (see Figure 3.1.20).
3. For every planet, the ratio of the cube of the average orbital radius, r, to the square of the period, T, of revolution is the same, i.e. $\frac{r^3}{T^2}$ = a constant, k.

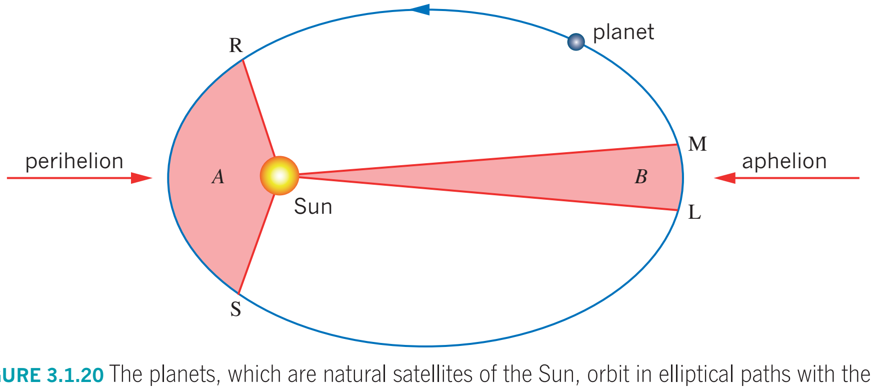


FIGURE 3.1.20 The planets, which are natural satellites of the Sun, orbit in elliptical paths with the Sun at one focus. Their speeds vary continually, and they are fastest when closest to the Sun. A line joining a planet to the Sun will sweep out equal areas in equal times. So, for example, the time it takes to move from R to S is equal to the time it takes to move from L to M, and so area *A* is the same as area *B*.

Kepler's first two laws proposed that planets moved in elliptical paths from furthest point (the *aphelion*) to closest point (the *perihelion*). The closer the planet was to the Sun, the faster it moved. It took Kepler many months of laborious calculations to arrive at his third law. Newton used Kepler's laws to justify the inverse square relationship. In fact, Kepler's third law can be deduced, for circular orbits, from Newton's law of universal gravitation.

CALCULATING THE ORBITAL PROPERTIES OF SATELLITES

The speed, v, of a satellite can be calculated from its motion for one revolution. It will travel a distance equal to the circumference of the circular orbit, $2\pi r$, in the time of one period, T.

The speed, v, of a satellite in a circular orbit is given by:

$$v = \frac{\text{distance}}{\text{time}} = \frac{2\pi r}{T}$$

where r is the radius of the orbit (m)

T is the time for one revolution, or the period (s)

The centripetal acceleration of a satellite can be determined from the gravitational field strength at its location. Satellites are in free fall; therefore, the only force acting is gravity, F_g. The International Space Station (ISS) is in orbit at a distance from Earth where g is 8.8 N kg^{-1}, and so it orbits with a centripetal acceleration of 8.8 m s^{-2}.

The centripetal acceleration, a, of the satellite can also be calculated by considering its circular motion. The equation for speed given above can be substituted into the centripetal acceleration formula to give:

$$a = \frac{v^2}{r} \text{ and since } v = \frac{2\pi r}{T}$$

$$\text{then } \frac{v^2}{r} = \frac{4\pi^2 r}{T^2}$$

Since the centripetal acceleration of the satellite is equal to the gravitational field strength at the location of its orbit, and using the gravitational field strength equation from Chapter 1, we can give the following expression.

The centripetal acceleration, a, of a satellite in circular orbit is given by:

$$a = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2} = \frac{GM}{r^2} = g$$

where v is the speed of the satellite (m s^{-1})

r is the radius of the orbit (m)

T is the period of orbit (s)

M is the central mass (kg)

g is the gravitational field strength at r (N kg^{-1})

G is the gravitational constant, 6.67×10^{-11} N m^2 kg^{-2}

These relationships can be manipulated to determine any feature of a satellite's motion: its speed, radius of orbit or period of orbit. They can also be used to find the mass of the central body around which the satellite orbits, M.

In the same way as with freely falling objects at the Earth's surface, the mass of the satellite itself has *no effect* on any of these orbital properties.

The gravitational force, F_g, acting on the satellite can then be found by using Newton's second law.

The gravitational force on a satellite of mass m in a stable circular orbit is given by:

$$F_g = \frac{mv^2}{r} = \frac{4\pi^2 rm}{T^2} = \frac{GMm}{r^2} = mg$$

Worked example 3.1.2

WORKING WITH KEPLER'S LAWS

Determine the orbital speed of the Moon, assuming it is in a circular orbit of radius 384 000 km around the Earth. Take the mass of the Earth to be 5.97×10^{24} kg and use $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2}.

Thinking	Working
Ensure that the variables are in their standard units.	$r = 384\,000$ km $= 3.84 \times 10^8$ m
Choose the appropriate relationship between the orbital speed, v, and the data that has been provided.	$a = g = \frac{GM}{r^2} = \frac{v^2}{r}$
Make v, the orbital speed, the subject of the equation.	$v = \sqrt{\frac{GM}{r}}$
Substitute in values and solve for the orbital speed, v.	$v = \sqrt{\frac{GM}{r}}$ $= \sqrt{\frac{(6.67 \times 10^{-11}) \times (5.97 \times 10^{24})}{3.84 \times 10^8}}$ $= 1.02 \times 10^3$ m s^{-1}

Worked example: Try yourself 3.1.2

WORKING WITH KEPLER'S LAWS

Determine the orbital speed of a satellite, assuming it is in a circular orbit of radius of 42 100 km around the Earth. Take the mass of the Earth to be 5.97×10^{24} kg and use $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2}.

HOW NEWTON DERIVED KEPLER'S THIRD LAW USING ALGEBRA

It took Kepler many months of trial-and-error calculations to arrive at his third law:

$$\frac{r^3}{T^2} = \text{constant.}$$

Newton was able to use some clever algebra to derive this from his law of universal gravitation:

$$F_g = \frac{m4\pi^2 r}{T^2} = \frac{GMm}{r^2} = mg$$

$$\therefore \frac{r^3}{T^2} = \frac{GM}{4\pi^2}$$

For any central mass, M, the term $\frac{GM}{4\pi^2}$ is constant and the ratio $\frac{r^3}{T^2}$ is equal to this constant value for all of its satellites (see Figure 3.1.21).

So, for example, if you know the orbital radius, r, and period, T, of one of the moons of Saturn, you could calculate $\frac{r^3}{T^2}$ and use this as a constant value for all of Saturn's moons. If you knew the period, T, of a different satellite of Saturn, it would then be straightforward to calculate its orbital radius, r.

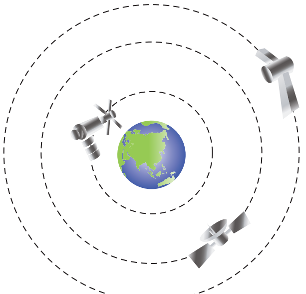


FIGURE 3.1.21 These three satellites are at different distances from Earth and hence according to Kepler's third law will have different orbital periods. For all three, the ratio of $\frac{r^3}{T^2}$ will equal the same constant value.

PHYSICSFILE

Ganymede

Jupiter is orbited by more than 60 known satellites, the biggest of which is Ganymede. Ganymede is very large. It is the biggest of all the moons in the solar system and is even bigger than the planet Mercury.

FIGURE 3.1.22 Ganymede.

Worked example 3.1.3

SATELLITES IN ORBIT

Ganymede is the largest of Jupiter's moons. It has a mass of 1.66×10^{23} kg, an orbital radius of 1.07×10^{6} km and an orbital period of 6.18×10^{5} s (7.15 days).

a Use Kepler's third law to calculate the orbital radius (in km) of Europa, another moon of Jupiter, which has an orbital period of 3.55 days.

Thinking	Working
Note down the values for the known satellite. You can work in days and km as this question involves ratio.	Ganymede: $r = 1.07 \times 10^{6}$ km $T = 7.15$ days
For all satellites of a central mass, $\frac{r^3}{T^2}$ = constant. Work out this ratio for the known satellite.	$\frac{r^3}{T^2}$ = constant $= \frac{(1.07 \times 10^{6})^3}{7.15^2}$ $= 2.40 \times 10^{16}$
Use this constant value with the ratio for the satellite in question. Make sure T is in days to match the ratio calculated in the previous step.	Europa: $T = 3.55$ days, r = ? $\frac{r^3}{T^2}$ = constant $\frac{r^3}{3.55^2} = 2.40 \times 10^{16}$
Make r^3 the subject of the equation.	$r^3 = 2.40 \times 10^{16} \times 3.55^2$ $= 3.02 \times 10^{17}$
Solve for r. The unit for r is km as the original ratio was calculated using km.	$r = \sqrt[3]{3.02 \times 10^{17}}$ $= 6.71 \times 10^{5}$ km Note: Europa has a shorter period than Ganymede so you should expect Europa to have a smaller orbit than Ganymede.

b Use the orbital data for Ganymede to calculate the mass of Jupiter.	
Thinking	**Working**
Note down the values for the known satellite. You must work in SI units to find the mass value in kg.	Ganymede/Jupiter: $r = 1.07 \times 10^9$ m $T = 6.18 \times 10^5$ s $m = 1.66 \times 10^{23}$ kg $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2} $M = ?$
Select the expressions from the equation for centripetal acceleration that best suit your data. $a = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2} = \frac{GM}{r^2} = g$	Use the 3rd and 4th terms of the expression. $\frac{4\pi^2 r}{T^2} = \frac{GM}{r^2}$ These two expressions use the given variables r and T, and the constant G, so that a solution may be found for M.
Transpose to make M the subject.	$M = \frac{4\pi^2 r^3}{GT^2}$
Substitute values and solve.	$M = \frac{4\pi^2(1.07 \times 10^9)^3}{6.67 \times 10^{-11} \times (6.18 \times 10^5)^2}$ $= 1.90 \times 10^{27}$ kg

c Calculate the orbital speed of Ganymede in km s^{-1}.	
Thinking	**Working**
Note values you will need to use in the equation $v = \frac{2\pi r}{T}$.	Ganymede: $r = 1.07 \times 10^6$ km $T = 6.18 \times 10^5$ s $v = ?$
Substitute values and solve. The answer will be in km s^{-1} if r is expressed in km.	$v = \frac{2\pi r}{T}$ $= \frac{2\pi \times 1.07 \times 10^6}{6.18 \times 10^5}$ $= 10.9$ km s^{-1}

Worked example: Try yourself 3.1.3

SATELLITES IN ORBIT

Callisto is the second largest of Jupiter's moons. It is about the same size as the planet Mercury. Callisto has a mass of 1.08×10^{23} kg, an orbital radius of 1.88×10^6 km and an orbital period of 1.44×10^6 s (16.7 days).

a Use Kepler's third law to calculate the orbital radius (in km) of Europa, another moon of Jupiter, which has an orbital period of 3.55 days.

b Use the orbital data for Callisto to calculate the mass of Jupiter.

c Calculate the orbital speed of Callisto in km s^{-1}.

3.1 Review

SUMMARY

- A normal force, F_N, is the force that a surface exerts on an object that is in contact with it. It acts at right angles to the surface and changes as the force exerted on the surface changes.
- The apparent weight of an object is equal to the normal reaction force acting on the object.
- Apparent weight increases or decreases as the surface you are standing on accelerates up or down.
- Astronauts in orbit experience apparent weightlessness as they are in free fall around the Earth.
- A satellite is an object that is in a stable orbit around a larger central mass.
- The only force acting on a satellite is the gravitational attraction between it and the central body.
- Satellites are in continual free fall. They move with a centripetal acceleration that is equal to the gravitational field strength at the location of their orbit.
- The speed of a satellite, v, is given by:
  $v = \frac{2\pi r}{T}$
- For a satellite in a circular orbit:
  $a = \frac{v^2}{r} = \frac{4\pi^2 r}{T} = \frac{GM}{r^2} = g$
- The gravitational force acting on a satellite in a circular orbit is given by:
  $F_g = \frac{mv^2}{r} = \frac{4\pi^2 rm}{T^2} = \frac{GMm}{r^2} = mg$
- For any central body of mass, M:
  $\frac{r^3}{T^2} = \frac{GM}{4\pi^2}$ = constant, so knowing another satellite's orbital radius, r, enables its period, T, to be determined.

KEY QUESTIONS

1 Determine the weight of a 6.50 kg box at the surface of the Earth where $g = 9.80 \text{ m s}^{-2}$ downwards.

2 A box of weight 150 N sits at rest on the floor. What is the magnitude of the normal force acting on the box?

3 Calculate the apparent weight of a 45.0 kg child standing in a lift that is accelerating upwards at 2.02 m s^{-2}.

4 Calculate the apparent weight of a 45.0 kg child standing in a lift that is moving upwards at a constant speed of 4.04 m s^{-1}.

5 Which of the following objects has the greatest apparent weight?
A a fly flying horizontally
B a walking fly
C a show-jumping horse mid jump
D the International Space Station

6 Which of the following is correct?
A Earth is a satellite of Mars.
B The Moon is a satellite of the Sun.
C Earth is a satellite of the Sun.
D The Sun is a satellite of Earth.

7 A geostationary satellite orbits above Singapore, which is on the equator. Which of the following statements about the satellite is correct?
A It is in a low orbit.
B It is in a high orbit.
C It passes over the north pole.
D It is not moving.

8 A satellite of mass M is in a circular orbit around the Moon. A module of mass M then attaches to the original satellite so that the combined mass is now $2M$. How does this affect the orbital properties of the satellite?
A The speed of the satellite will decrease and the period will increase.
B Both the speed and period of the satellite will decrease.
C Both the speed and period of the satellite will increase.
D Nothing will change.

9 The gravitational field strength at the location where the Optus D1 satellite is in stable orbit around the Earth is equal to 0.22 N kg^{-1}. The mass of this satellite is 2.3×10^3 kg.
a Using only the information given, calculate the magnitude of the acceleration of this satellite as it orbits.
b Calculate the net force acting on this satellite as it orbits.

10 One of Saturn's moons is Atlas, which has an orbital radius of 1.37×10^5 km and a period of 0.60 days. The largest of Saturn's moons is Titan. It has an orbital radius of 1.20×10^6 km. What is the orbital period of Titan in days?

3.2 DC motors

Physicists have always been interested in the relationship between electricity and magnetism because they wanted to understand the basic workings of the universe. For the world at large, however, this understanding provided a more practical form of excitement. It enabled the generation and use of electricity on a large scale. One of the most obvious applications of the understanding of electromagnetism gained in the 19th century is the electric motor.

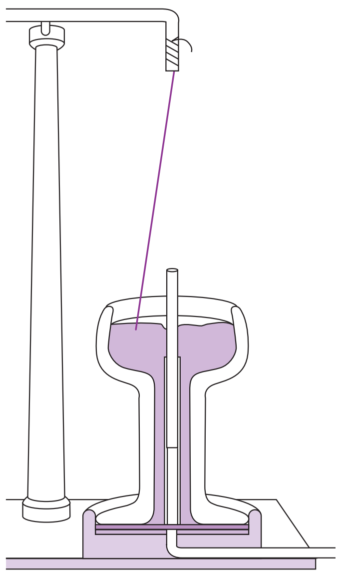


FIGURE 3.2.1 Michael Faraday's electric motor.

DC MOTORS

The main components and the principles have been the same for all DC motors since Michael Faraday built the first one in 1821 (see Figure 3.2.1). In Faraday's motor, a magnet was mounted vertically in a pool of mercury. A wire carrying a current hung from a support above. (The mercury provided a path for the current.) The magnetic field of the magnet spread outwards from the top of the magnet and so there was a component of this field that was perpendicular to the wire. This produced a horizontal force on the wire that kept it rotating around the magnet. Use the right-hand rule from the previous chapter to convince yourself that if the current flows down and the magnetic field points out from the central magnet, the wire will rotate clockwise when viewed from above.

In modern **direct current** (DC) motors, a current-carrying coil of wire in a magnetic field experiences a magnetic force, F, equal to $nIlB$ on two or more of its sides. In practice, many turns of wire (n) are used and the magnetic field is provided by more than one permanent magnet or by an **electromagnet**.

The formula $F = nIlB$ includes the number of coils of wire, n, which equals 1 for all the examples in this section. Therefore, $F = IlB$ will be used to solve problems throughout this section.

Consider a single square coil of wire, with vertices ABCD, carrying a current, I, in a magnetic field, B, as shown in Figure 3.2.2.

Initially the wire coil is aligned horizontally in a magnetic field, B, as in Figure 3.2.2(a). Sides AD and BC are parallel to the magnetic field so no magnetic force will act on them. Sides AB and CD are perpendicular to the field so both of these sides will experience a magnetic force. Using the right-hand rule, there is a downwards force on AB and an upwards force on CD. These two forces will act together on the coil and cause it to rotate anticlockwise. If the coil is free to turn it will move toward the position shown in Figure 3.2.2(b).

In Figure 3.2.2(b), there will be a magnetic force acting on every side of the coil. However, the forces acting on sides AD and BC will be equal and opposite in direction. They will tend to stretch the coil outwards but won't affect its rotation. The forces on sides AB and CD will remain and the coil will continue to rotate anticlockwise.

As the coil rotates to the position shown in Figure 3.2.2(c), the forces acting on each side are such that they will tend to keep the coil in this position. The force on each side will act outwards from the coil. There are no turning forces at this point, but any further rotation will cause a force in the opposite direction that will cause the coil to rotate clockwise, back to this perpendicular position. For the coil to continue to rotate anticlockwise at this point, the current direction needs to be reversed. This is shown in Figure 3.2.2(d). With the current reversed, all of the forces are reversed, and provided the coil has a little momentum to get it past the perpendicular position, it will continue to rotate anticlockwise. This ability to reverse the current direction at the point where the coil is perpendicular to the magnetic field is a key design feature in DC motors. It is a **commutator** that allows the current to be reversed.

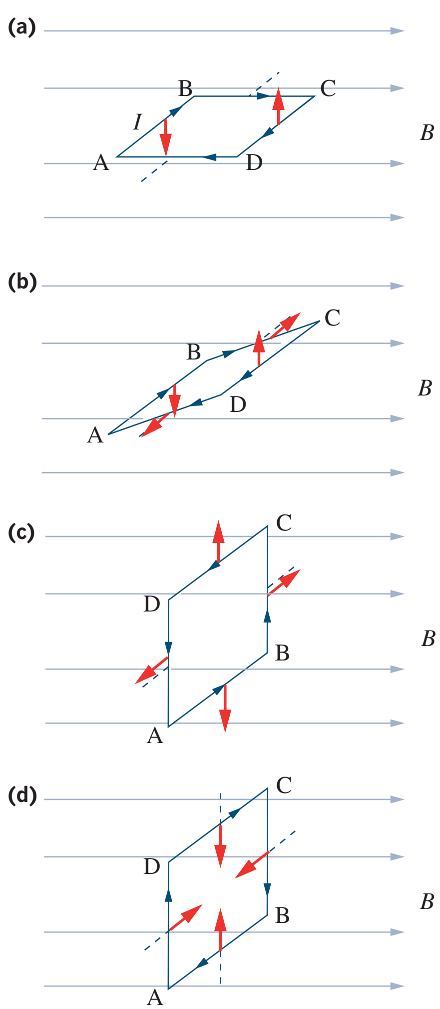


FIGURE 3.2.2 The magnetic force acting on each side of a current-carrying square wire coil in a magnetic field, B.

PHYSICSFILE

Michael Faraday

Michael Faraday (1791–1867), depicted in Figure 3.2.3, was an English scientist who worked in the areas of chemistry and physics. He had little formal education. At the age of fourteen he became the apprentice to a London bookbinder. During his apprenticeship he read many of the books that came his way. At the age of 21 he became a laboratory assistant to Sir Humphry Davy, who was one of the most prominent scientists of the day. Faraday was a gifted experimenter and after returning from a scientific tour through Europe with Davy, he began to be recognised in his own right for the scientific work he was doing. He was admitted to the Royal Society at age 32. He is credited with the discoveries of benzene, electromagnetic induction and the basis of the modern electric motor. He died in 1867 at Hampton Court. His contributions to science, and in particular his work in the area of electromagnetism, are recognised through the unit of measurement of capacitance known as the farad. In Chapter 4 'Electromagnetic induction and transmission of electricity', you will study more of Michael Faraday's work on electromagnetic induction.

FIGURE 3.2.3 Michael Faraday.

Torque

The turning force that the coil experiences in an electric motor is referred to as the **torque** on the coil. A torque is the turning effect of any force, for example, pushing on a swinging door. To achieve the maximum effect, the force should be applied at right angles to the door and at the largest distance possible from the point where the door is hinged. This idea is illustrated in Figure 3.2.4.

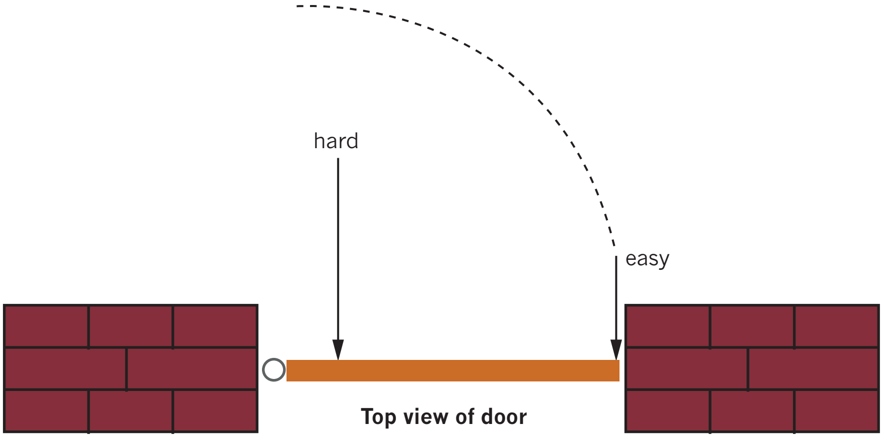


FIGURE 3.2.4 The force required to open a swinging door decreases as the perpendicular distance from the point of rotation increases and the torque, or turning effect, is maximised.

Torque is thus defined as:

$\tau = r_{\perp}F$

where τ is the torque (N m)

$r_{\perp}$ is the perpendicular distance between the axis of rotation and the point of application of the force (m)

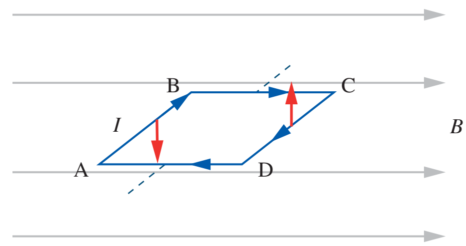
F is the component of the force perpendicular to the axis of rotation (N)

In the case of a single square or rectangular coil, each of the two sides perpendicular to the magnetic field will experience a force contributing to the total torque, hence the total torque applied to the coil will be twice that acting on one side.

Worked example 3.2.1

TORQUE ON A COIL

A single square wire coil, ABCD, of side length 5.00 cm, is free to rotate within a magnetic field, B, of strength 1.00×10^{-4} T. A current of 1.00 A is flowing through the coil. What is the torque on the coil?

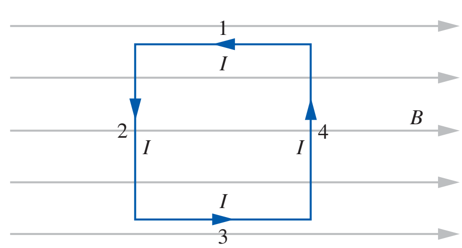


Thinking	Working
Confirm that the coil will experience a force based on the magnetic field and current directions supplied.	Using the right-hand rule confirms that a downwards force applies on side AB. An upwards force applies on side CD. The coil will turn anticlockwise. Sides AD and BC lie parallel to the magnetic field and no force will apply.
Calculate the magnetic force on one side.	$F = IlB$ $= 1.0 \times 0.0500 \times 1.00 \times 10^{-4}$ $= 5.00 \times 10^{-6}$ N
Determine the distance, r, from the point of rotation that the magnetic force is applied.	length of side = 5.00 cm distance between axis of rotation and application of force = $\frac{1}{2}$ × side length $r = 2.50$ cm $= 0.0250$ m
Calculate the torque applied by the magnetic force on one side of the coil.	$\tau = r_{\perp}F$ $= 0.0250 \times 5.00 \times 10^{-6}$ $= 1.25 \times 10^{-7}$ N m
Since two sides, AB and CD, both experience a magnetic force and hence a torque, the torque on one side should be multiplied by 2 to find the total torque. State the direction of rotation.	total torque = $2 \times 1.25 \times 10^{-7}$ $= 2.50 \times 10^{-7}$ N m The direction is anticlockwise.

Worked example: Try yourself 3.2.1

TORQUE ON A COIL

A single square wire coil, with a side length of 4.0 cm, is free to rotate within a magnetic field, B, of strength 1.0×10^{-4} T. A current of 1.0 A is flowing through the coil. What is the torque on the coil?



PRACTICAL DC MOTORS

A basic single-coil electric motor with a simple arrangement to reverse the current direction will work, but it won't turn very smoothly. That's because maximum torque will only apply each half turn or twice for every full rotation. A number of enhancements have been developed over time to make DC motors the highly practical motive force they are today.

The commutator is usually made from a split ring of copper or another good conductor on which conducting brushes (usually carbon blocks) rub. Each half is connected to one end of the coil of wire. This arrangement of brushes prevents the wire from becoming tangled as the coil rotates. The commutator reverses the current at the point where the coil is perpendicular to the magnetic field, which keeps the coil rotating (see Figure 3.2.5).

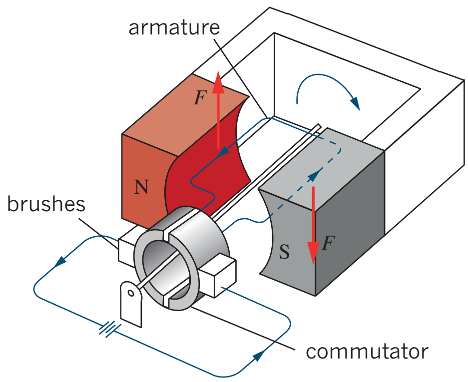


FIGURE 3.2.5 The main parts of a simple but practical single-coil DC electric motor.

Practical motors will have many sets of coils of many turns each, spaced at an angle to each other, as shown in Figure 3.2.6.

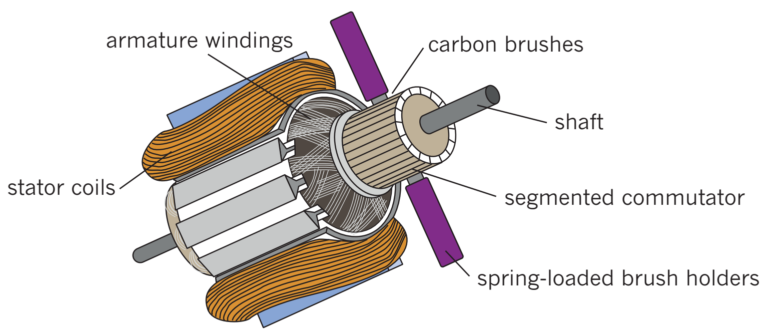


FIGURE 3.2.6 A typical multi-coil DC electric motor, showing the main components. Note that there are many sets of coils offset by an angle from each other. The stator coils produce an electromagnet that provides the magnetic field. The commutator feeds current to the armature coils in the position where maximum torque will be experienced.

The coils are wound around a soft iron core to increase the magnetic field that passes through them. The whole arrangement of core and coils is called an **armature** (as shown in Figure 3.2.6). Permanent magnets are generally used to provide the magnetic field in small motors, but in larger motors electromagnets are used as they can produce larger and stronger fields. These magnets are usually stationary, as distinct from the rotating rotor or armature, and are often referred to as the **stator**. The commutator is arranged to feed current to the particular coil that is in the best position to provide maximum torque. The total torque will be the sum of the torques on all the individual coils.

Generally speaking, the larger the torque in an electric motor the better. This is achieved by the use of a strong magnetic field, a large number of turns of wire in each coil, a high current and a large area of coil. All this adds to the cost, so when designing an electric motor, each aspect may be compromised to some extent in light of its potential use.

3.2 Review

SUMMARY

- The magnetic force on a current-carrying wire within a magnetic field is $F = nIlB$.
- There is a torque on a coil of wire carrying a current whenever the current is not parallel to the field. Torque is defined as: $\tau = r_{\perp}F$
- The wire coil of a simple DC motor keeps rotating because the direction of current, and hence torque, is reversed each half turn by the commutator.
- In the case of a single square or rectangular coil, the total torque applied to the coil will be twice that acting on the one side.
- The armature of a practical motor consists of many coils that are fed current by the commutator when they are in the position of maximum torque.
- The total torque will be the sum of the torques on all the individual coils.

KEY QUESTIONS

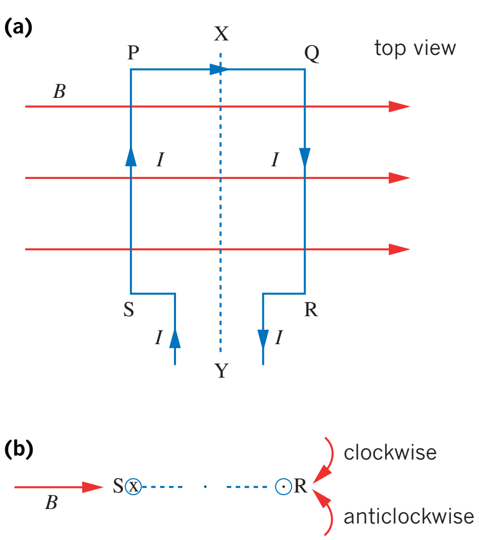
1 For which of the following situations is torque at a maximum?
  A when the force is applied perpendicular to the axis of rotation
  B when the force is applied parallel to the axis of rotation
  C when the force is applied at a maximum regardless of direction
  D when the force applied is zero

The information below applies to questions 2–7.

Part (a) of the diagram below depicts a top view of a single current-carrying coil in an external magnetic field *B*.

Part (b) of the diagram is the corresponding cross-sectional view as seen from point *Y*. The following data apply:

$B = 0.10$ T, PQ = 2.0 cm, PS = QR = 5.0 cm, $I = 2.0$ A.



2 What is the magnitude and direction of the magnetic force acting on side PS?

3 What is the magnitude and direction of the magnetic force acting on side QR?

4 What is the magnitude of the force on side PQ?

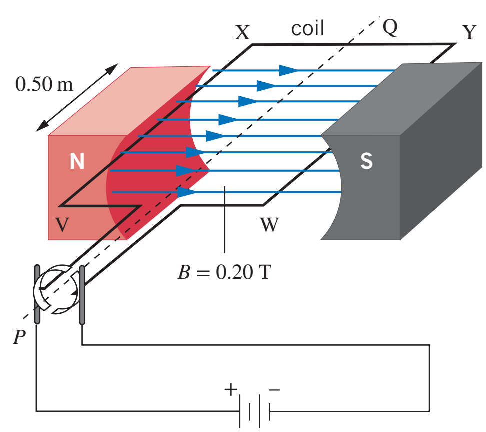
5 The coil is free to rotate about an axis through XY. In what direction, as seen from Y, would the coil rotate?

6 Which of the following does not affect the magnitude of the torque acting on this coil?
  A the dimensions of the coil
  B the magnetic field strength
  C the magnitude of the current through the coil
  D the direction of the current through the coil

7 What is the total torque acting on the coil?

The following information applies to questions 8–10.

The diagram shows a simplified version of a direct-current motor.



8 For the position of the coil shown, calculate the magnitude of the force on segment WY when a current of 1.0 A flows through the coil.

9 In which direction will the coil begin to rotate? Give your reasoning.

10 Which of the following actions would cause the coil to rotate faster?
  A increasing the current
  B increasing the magnetic field strength
  C increasing the cross-sectional area of the coil
  D all of the above

3.3 Particle accelerators

Melbourne is the home to the most powerful synchrotron in the southern hemisphere (see Figure 3.3.1). Looking something like a giant doughnut about 200 m in circumference, it produces beams of electromagnetic radiation, from infrared, through visible light, to 'hard' X-rays.

FIGURE 3.3.1 View of the inside of the Australian Synchrotron, taken from the mezzanine.

A **synchrotron** is a type of **particle accelerator**. Bunches of electrons are accelerated around a huge evacuated ring to almost the speed of light to energies as high as 3 billion electron-volts (3×10^9 eV). These charges are forced to follow a curved path, due to the magnetic field generated by bending magnets. As they accelerate around curves, the electrons give off bursts of radiation. This synchrotron radiation is channelled down tubes called beamlines and utilised by researchers in a range of experimental stations.

This section looks at the acceleration of charged particles in uniform electric and magnetic fields, including the change of speed caused by electric fields and the change of direction caused by magnetic fields.

PARTICLE ACCELERATORS

Particle accelerators are machines that were originally designed to investigate the nature of matter by examining the structure of atoms and molecules. Charged particles, such as electrons, protons or atomic nuclei, are accelerated to speeds often close to that of light. These particles travel through an electric field, inside a hollow tube pumped to an ultra-high vacuum, with pressures comparable to those found in deep space. Strong magnets direct the particles to collide with a target or with another moving particle. Scientists obtain information about the make-up of the subatomic particles fired from the machine, or the target samples that are hit, by analysing the types of collisions that occur.

One of the first particle accelerators was the Van de Graaff accelerator, similar to the Van de Graaff generator (see Figure 3.3.2). Developed in the 1930s, it can accelerate charged particles between metal electrodes to energies of about 15 MeV before they collide with a fixed target. Currently, the world's most powerful particle accelerator is the Large Hadron Collider. It is located at CERN on the France–Switzerland border. It can produce energies of 13 TeV. Two sets of particles can be accelerated in opposite directions around its central evacuated ring, to meet in a collision of mammoth energies!

FIGURE 3.3.2 This tandem Van de Graaff accelerator uses two generators to produce beams of charged particles that are accelerated by potential differences of up to 10 million volts.

ACCELERATING CHARGED PARTICLES: CHANGING THE SPEED

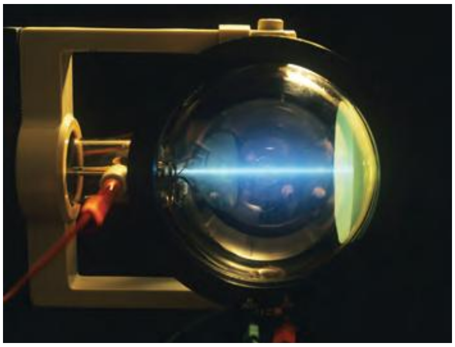


FIGURE 3.3.3 Cathode ray tube.

A **cathode ray tube** is a useful type of particle accelerator. Electrons are released from a negative terminal, or hot cathode, in a vacuum, and accelerate towards a positive terminal, or anode. The beam of electrons is collimated (narrowed) as it passes through a slit, and releases light when it hits a fluorescent screen. A potential difference of around 2–3 kV exists between the cathode and the anode, which causes the charged particles to accelerate. Older style televisions (before plasma, LCD and LED screens were invented), visual display units and cathode ray oscilloscopes (CROs) all consist of cathode ray tubes (see Figure 3.3.3).

A computer monitor, cathode ray oscilloscope or larger-scale particle accelerator relies on a source of charged particles to be accelerated. The device used to provide these particles is called an **electron gun**.

In an electron gun, electrons are, in effect, boiled off a heated wire filament, or cathode, shown on the left in Figure 3.3.4. They are accelerated from rest across an evacuated chamber towards a positively charged plate, or anode, due to the electric field created between charged plates (see Figure 3.3.4). Once the electrons continue through a gap in this positive plate, their motion can be further controlled by additional electric and magnetic fields. Focusing magnets are also used to control the width of the beam.

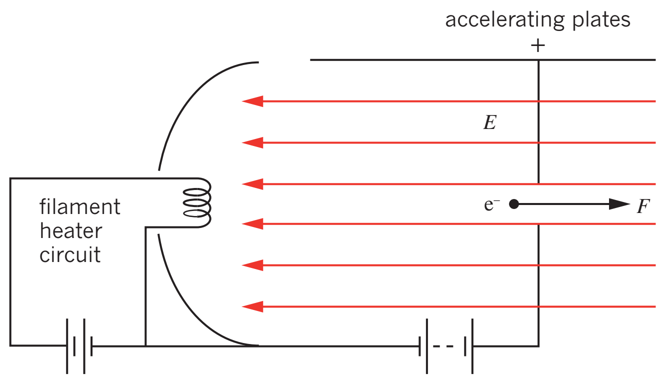


FIGURE 3.3.4 Electron-gun assembly.

Consider an electric field acting on an electron as the result of a pair of oppositely charged parallel plates connected to a DC power supply. The electron is attracted to the positive plate and repelled from the negative plate. An electric field is acting upon any charged particle within this region. This electric field is a vector quantity and may be compared in some ways to the Earth's gravitational field. Recall from Chapter 2 that an electric field has units N C^{-1} and is defined as:

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

where F is the force (N) experienced by a charged particle due to an electric field and q is the magnitude of the electric charge of a particle in the field, in this case an electron (1.6×10^{-19} C).

A charge will then experience a force equal to qE when placed within such an electric field.

Recall that the magnitude of the electric field may also be expressed as:

$$E = \frac{V}{d}$$

where d is the separation of the plates (m) and V is the potential difference (V). Combining these two relationships produces an expression for the force on a charge within a pair of parallel charged plates:

$$\frac{F}{q} = \frac{V}{d}$$

$$F = \frac{qV}{d}$$

In addition, calculations of the energy gained by an electron as it is accelerated towards a charged plate by the electric field can be made. The work done in this case is equivalent to:

$W = qV$

This equation can be used to calculate the increase in kinetic energy as an electron accelerates from one plate to another.

If a charge is accelerated from rest from an electron gun, then:

$E_k = W = qV$

$E_k = \frac{1}{2}mv^2 - \frac{1}{2}mu^2$

where v is the final velocity and u is the initial velocity of the charge. If the electron accelerates from rest ($u = 0$), then this can be simplified to:

$E_k = \frac{1}{2}mv^2 = qV$

> $\frac{1}{2}mv^2 = qV$
>
> This is often referred to as the electron-gun equation.

Worked example 3.3.1

CALCULATING THE SPEED OF ACCELERATED CHARGED PARTICLES

Determine the final speed of a single electron, with a charge of magnitude 1.6×10^{-19} C and a mass of 9.1×10^{-31} kg, when accelerating across a potential difference of 1.5 kV.	
Thinking	**Working**
Ensure that the variables are in their standard units.	$1.5 \text{ kV} = 1.5 \times 10^3 \text{ V}$
Establish what quantities are known and what are required.	$v = ?$ $q = 1.6 \times 10^{-19}$ C $m = 9.1 \times 10^{-31}$ kg $V = 1.5 \times 10^3$ V
Substitute values into the electron-gun equation and rearrange to solve for the speed.	$qV = \frac{1}{2}mv^2$ $1.6 \times 10^{-19} \times 1.5 \times 10^3 = \frac{1}{2} \times 9.1 \times 10^{-31} \times v^2$ $v = \sqrt{\frac{2 \times 1.6 \times 10^{-19} \times 1.5 \times 10^3}{9.1 \times 10^{-31}}}$ $= 2.3 \times 10^7 \text{ m s}^{-1}$

Worked example: Try yourself 3.3.1

CALCULATING THE SPEED OF ACCELERATED CHARGED PARTICLES

> Determine the final speed of a single electron, with a charge of magnitude 1.6×10^{-19} C and a mass of 9.1×10^{-31} kg, when accelerating across a potential difference of 1.2 kV.

THE EFFECT ON A CHARGED PARTICLE IN A MAGNETIC FIELD

To explore the forces acting on a beam of electrons in a particle accelerator (see Figure 3.3.5), the effect of a magnetic field on a charged particle also needs to be considered. From Chapter 2, recall that because an electric current is itself a stream of moving charges, the magnitude of the force, F, on a charge, q, moving with velocity, v, perpendicular to a magnetic field of strength B is given by:

$F = qvB$

In the case of the magnetic force on an electron moving within the magnetic field of a particle accelerator, the magnitude of charge, q, is equal to 1.6×10^{-19} C.

The direction of the magnetic force exerted on the charge is predicted by the right-hand rule. Note that the direction of current is defined as the direction in which a positive charge would move, so this direction must be reversed to correctly predict the direction of motion of an electron.

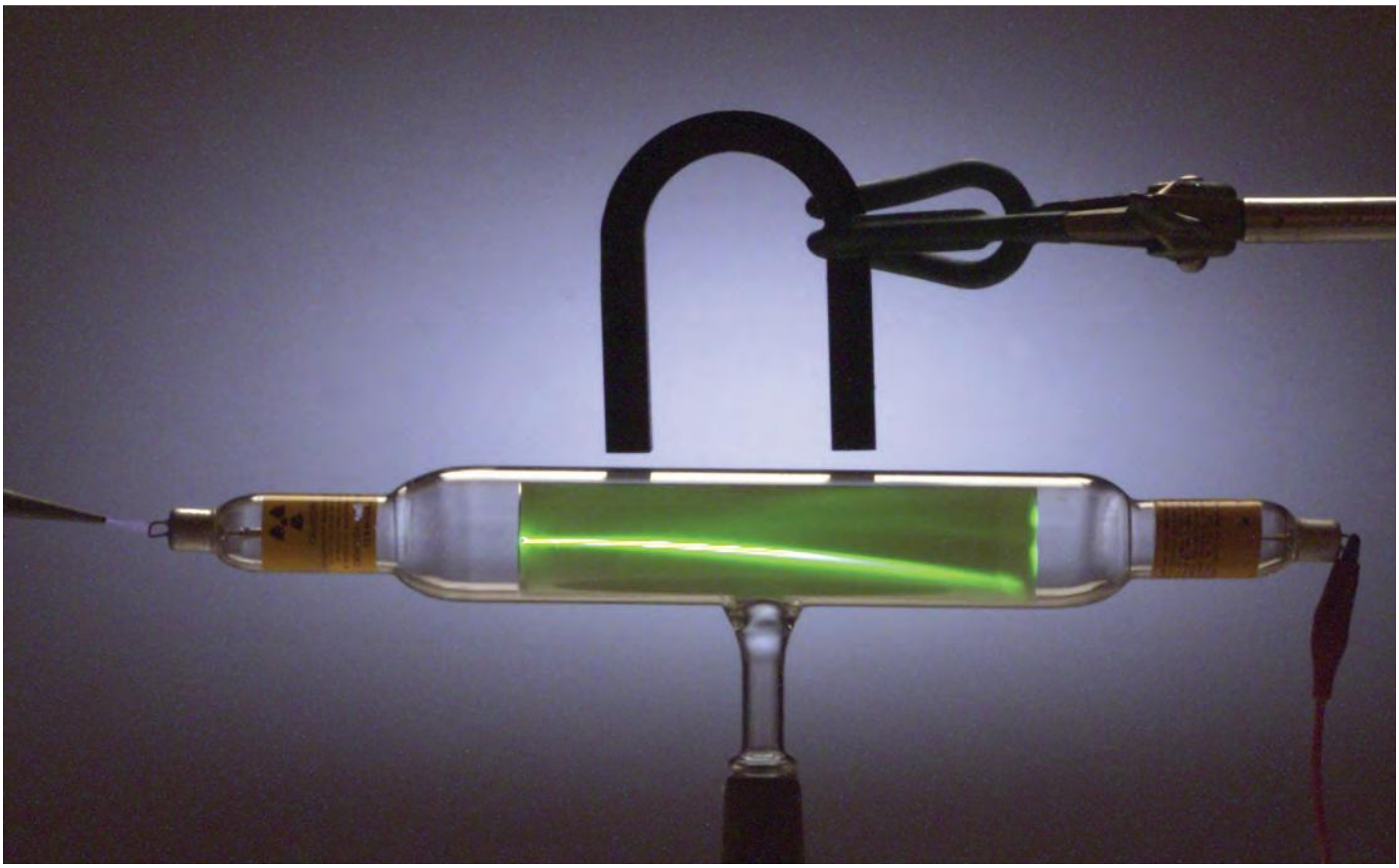


FIGURE 3.3.5 Electron beam being deflected by a magnet.

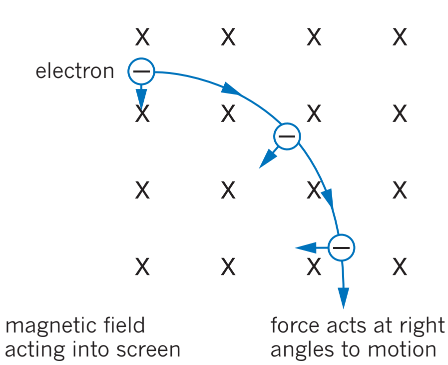


FIGURE 3.3.6 An electron moving in a magnetic field.

If a moving charge experiences a force of constant magnitude that remains at right angles to its motion, its direction will be changed but not its speed. In this way, bending magnets within a particle accelerator act to alter the path of the electron beam, rather than to speed the electrons up. As a result, the electrons will follow a curved path of radius r, as shown in Figure 3.3.6.

In this case, the net force acting on the charge is:

$F = ma$

This is equivalent to the magnetic force on the charge, so that:

$qvB = ma$

The acceleration in this situation is centripetal (towards the centre of the circular path) and has magnitude:

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

Substituting this relationship into the previous equation gives:

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

Rearranging this equation gives an expression that predicts the radius of the path of an electron travelling at right angles to a constant magnetic field:

$$r = \frac{mv}{qB}$$

where r is the radius of the path (m)

m is the mass of the electron (9.1×10^{-31} kg)

v is the speed of the charge (m s^{-1})

q is the charge on the electron (-1.6×10^{-19} C)

B is the strength of the magnetic field (T)

This relationship can be used to calculate the radius of the path followed by an electron travelling at right angles to any magnetic field. The electron could be a low-velocity electron or could be a high-velocity electron that has been accelerated by the powerful bending magnets within a particle accelerator.

Worked example 3.3.2

CALCULATING SPEED AND PATH RADIUS OF ACCELERATED CHARGED PARTICLES

An electron gun releases electrons from its cathode which are then accelerated across a potential difference of 32 kV, over a distance of 30 cm between a pair of charged parallel plates. Assume that the mass of an electron is 9.1×10^{-31} kg and the magnitude of the charge on an electron is 1.6×10^{-19} C.

a Calculate the strength of the electric field acting on the electron beam.

Thinking	Working
Ensure that the variables are in their standard units.	$32 \text{ kV} = 32 \times 10^3 = 3.2 \times 10^4$ V 30 cm = 0.30 m
Apply the correct equation.	$E = \frac{V}{d}$
Solve for E.	$E = \frac{3.2 \times 10^4}{0.30}$ $= 1.1 \times 10^5$ V m^{-1}

b Calculate the speed of the electrons as they exit the electron gun assembly.

Thinking	Working
Apply the correct equation.	$\frac{1}{2}mv^2 = qV$
Rearrange the equation to make v the subject.	$v = \sqrt{\frac{2qV}{m}}$
Solve for v.	$v = \sqrt{\frac{2 \times 1.6 \times 10^{-19} \times 3.2 \times 10^4}{9.1 \times 10^{-31}}}$ $= 1.1 \times 10^8$ m s^{-1}

c The electrons then travel through a uniform magnetic field perpendicular to their motion. Given that this field is of strength 0.2 T, calculate the expected radius of the path of the electron beam.	
Thinking	**Working**
Apply the correct equation.	$r = \frac{mv}{qB}$
Solve for *r*.	$r = \frac{9.1 \times 10^{-31} \times 1.1 \times 10^{8}}{1.6 \times 10^{-19} \times 0.2}$ $= 3.1 \times 10^{-3}$ m

Worked example: Try yourself 3.3.2

CALCULATING SPEED AND PATH RADIUS OF ACCELERATED CHARGED PARTICLES

An electron gun releases electrons from its cathode which are accelerated across a potential difference of 25 kV, over a distance of 20 cm between a pair of charged parallel plates. Assume that the mass of an electron is 9.1×10^{-31} kg and the magnitude of the charge on an electron is 1.6×10^{-19} C.

a Calculate the strength of the electric field acting on the electron beam.

b Calculate the speed of the electrons as they exit the electron-gun assembly.

c The electrons then travel through a uniform magnetic field perpendicular to their motion. Given that this field is of strength 0.3 T, calculate the expected radius of the path of the electron beam.

PHYSICS IN ACTION

Thomson's $\frac{e}{m}$ experiment

The knowledge and use of the properties of electrons are only relatively recent accomplishments in science. It was not until 1897 that physicists were able to shed any light on the internal physical structure of the atom. In that year, Joseph John Thomson demonstrated that cathode rays—rays emanating from a heated cathode in a vacuum—were particles that are fundamental constituents of every atom. For the first time, the atom was shown to have component particles rather than being indivisible. To indicate their importance, cathode rays were renamed electrons.

Thomson's experiment with cathode rays was performed in two stages (see Figure 3.3.7). At first the forces on a beam of electrons were balanced using an electric and a magnetic field, as shown by the central dotted line striking the fluorescent screen in Figure 3.3.7. This enabled Thomson to find the speed of the electrons. Then the magnetic field was switched off, and the beam was deflected under the influence of the electric field alone, as shown by the upper dotted line striking the fluorescent screen in Figure 3.3.7. The deflection of the beam was measured, allowing Thomson to find the charge-to-mass ratio $\left(\frac{e}{m}\right)$ for the cathode rays. Thomson repeated the experiment with a variety of different cathodes to show that all cathode rays yielded the same value. His result produced a value of about 1×10^{11} C kg^{-1}; the accepted value today is 1.76×10^{11} C kg^{-1}.

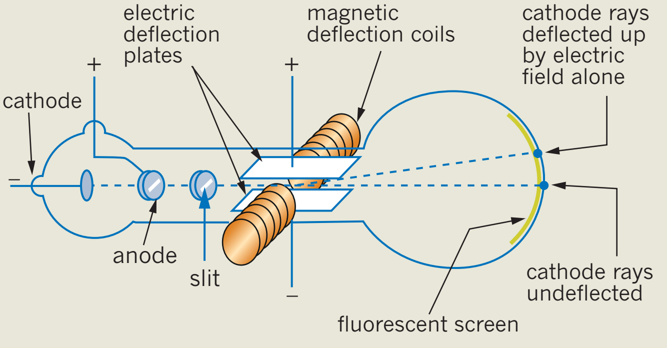


FIGURE 3.3.7 J. J. Thomson's apparatus for finding the charge to mass ratio $\left(\frac{e}{m}\right)$ for cathode rays (electrons). In 1897, Thomson used an electron gun to produce a beam of electrons that could be deflected by an electric and magnetic field within an evacuated tube.

PARTICLE ACCELERATORS

To study the basic constituents of matter, physicists accelerate particles such as electrons and protons to very high speeds before crashing them into other particles. The by-products from these collisions have revealed a vast array of sub particles and led to a better understanding of the fundamental properties of these particles. One of the key findings of recent particle accelerator experiments is the Higgs boson.

The particles are accelerated by electromagnetic fields, but very long paths are required for the particles to obtain the extremely high speeds needed (very close to the speed of light). To achieve this without the need for tunnels hundreds of kilometres long, particles travel through very strong magnetic fields that cause them to move in a circle. The Australian Synchrotron, near Monash University in Melbourne, is 70 m in diameter.

The Australian Synchrotron accelerates electrons through an equivalent of 3000 million volts (3 GV). At this energy, they travel at 99.99999% of the speed of light. Because of the relativistic effects that occur at these near-light speeds, their effective mass is about 6000 times that at rest. Because they are being accelerated, the electrons emit electromagnetic radiation. It is this light, ranging from infrared through to X-ray wavelengths, that is used for the research projects being conducted at the synchrotron (see Figure 3.3.8).

FIGURE 3.3.8 An inside view of the Australian Synchrotron.

3.3 Review

SUMMARY

- Particle accelerators are machines that accelerate charged particles, such as electrons, protons or atomic nuclei, to speeds close to that of light.
- The device used to provide these particles is called an electron gun.
- The work done on a charged particle in an electric field causes a change in the kinetic energy of the particle. If the particle is accelerated from rest, the work done is equal to the final kinetic energy, $W = qV = \frac{1}{2}mv^2$
- The magnitude of the force on a charged object within a magnetic field is given by $F = qvB$.
- The right-hand rule is used to determine the direction of the force on a positive charge moving in a magnetic field, B. The direction of the force on a negatively charged particle is in the opposite direction.
- The radius of the path of an electron travelling at right angles to a uniform magnetic field is given by $r = \frac{mv}{qB}$.

KEY QUESTIONS

1 How are particle accelerators able to provide the centripetal acceleration to change the direction of a charged particle using electromagnetic fields?

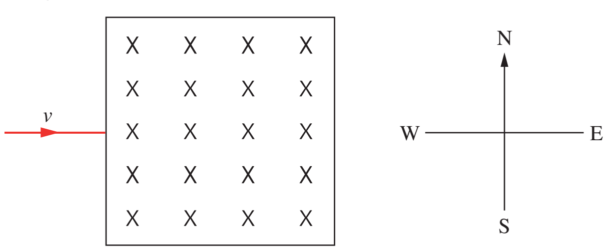
A Charged particles are part of the electromagnetic spectrum.

B Charged particles experience a force from the magnetic field that is proportional to the particle's velocity, constantly accelerating the charged particle.

C The accelerator is curved around the magnetic field.

D Charged particles will always accelerate when placed in a vacuum.

2 An electron with a charge magnitude of 1.6×10^{-19} C is moving eastwards into magnetic field of strength $B = 1.5 \times 10^{-5}$ T acting into the screen, as shown below. If the magnitude of the initial velocity is 1.0 m s^{-1}, what is the magnitude and direction of the force it initially experiences as it enters the magnetic field?



3 Electrons in a cathode ray tube (CRT) are accelerated through a potential difference of 2.5 kV. Calculate the speed at which they hit the screen of the CRT.

4 An electron travelling at a speed of 7.0×10^6 m s^{-1} passes through a magnetic field of strength 8.6×10^{-3} T. The electron moves at right angles to the field.

a Calculate the force exerted on the electron by the magnetic field.

b Given that this force directs the electron in a circular path, calculate the radius of its motion.

5 An electron with speed 7.6×10^6 m s^{-1} travels through a uniform magnetic field and follows a circular path of diameter 9.2×10^{-2} m. Calculate the magnetic field strength through which the electron travels.

6 In an experiment similar to Thomson's for determining the charge to mass ratio $\frac{e}{m}$ of cathode rays (electrons), electrons travel at right angles through a magnetic field of strength 1.5×10^{-4} T. Given that they travel in an arc of radius 6 cm and that $\frac{e}{m} = 1.76 \times 10^{11}$ C kg^{-1}, calculate the speed of the electrons.

7 A particle accelerator uses magnetic fields to accelerate electrons to very high speeds. Explain, using appropriate theory and relationships, how the accelerator achieves these high speeds.

8 An electron beam travelling through a cathode ray tube is subjected to simultaneous electric and magnetic fields. The electrons emerge with no deflection. Given that the potential difference across the parallel plates X and Y is 3.0 kV, and that the applied magnetic field is of strength 1.6×10^{-3} T, calculate the distance between the plates.

Chapter review

03

KEY TERMS

apparent weight
apparent weightlessness
armature
artificial satellites
cathode ray tube
centripetal acceleration
commutator
direct current
electromagnet
electron gun
free fall
geostationary satellite
natural satellite
normal reaction force
particle accelerator
satellite
stator
synchrotron
torque
weight

1 Calculate the apparent weight of a 45.0 kg child standing in a lift that is decelerating at 3.15 m s^{-2} while travelling upwards.

2 Which description best describes the motion of astronauts when orbiting the Earth?
- **A** They float in a zero gravity environment.
- **B** They float in a reduced gravity environment.
- **C** They fall down very slowly due to the very small gravity.
- **D** They fall in a reduced gravity environment.

3 Select the statement below that correctly states how a satellite in a stable circular orbit 200 km above the Earth will move.
- **A** It will have an acceleration of 9.8 m s^{-2}.
- **B** It will have constant velocity.
- **C** It will have zero acceleration.
- **D** It will have acceleration of less than 9.8 m s^{-2}.

4 What can be said about an object if that object is orbiting the Earth in space and appears to be weightless?
- **A** It is in free fall.
- **B** It is in zero gravity.
- **C** It has no mass.
- **D** It is floating.

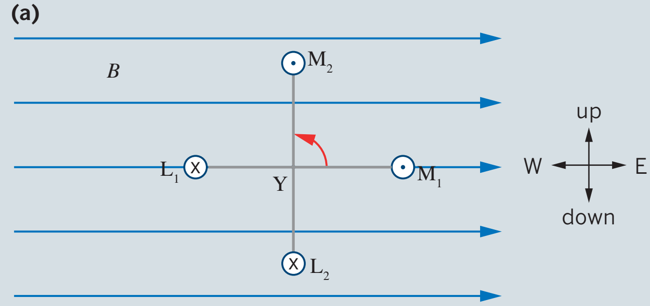
5 A low-Earth-orbit satellite X has an orbital radius of *r* and period *T*. A high-Earth-orbit satellite Y has orbital radius of 5*r*. In terms of *T*, what is the orbital period of Y?

6 The planet Neptune has a mass of 1.02×10^{26} kg. One of its moons, Triton, has a mass of 2.14×10^{22} kg and an orbital radius equal to 3.55×10^{8} m.
- **a** Calculate the orbital acceleration of Triton.
- **b** Calculate the orbital speed of Triton.
- **c** Calculate the orbital period of Triton (in days).

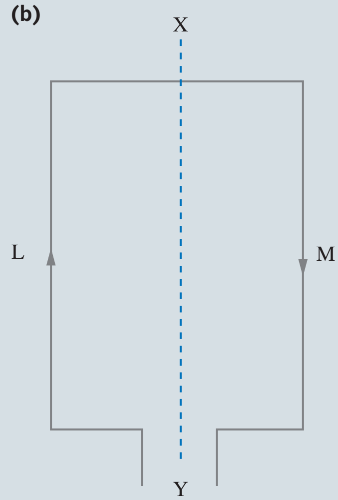
7 Ceres, the first asteroid to be discovered, was found by Giuseppe Piazzi in 1801. Ceres has a mass of 7.0×10^{20} kg and a radius of 385 km.
- **a** What is the gravitational field strength at the surface of Ceres?
- **b** Determine the speed required by a satellite in order to remain in orbit 10 km above the surface of Ceres.

The following information applies to questions 8–11.
Diagram (a) below shows an end-on view of a current-carrying loop, LM. The loop is free to rotate about a horizontal axis XY. You are looking at the loop from the Y end of the axis. The same loop is seen from the top in figure (b). Initially, arms L and M are horizontal (L_1–M_1). Later they are rotated so that they are vertical (L_2–M_2). The loop is located in an external magnetic field of magnitude *B* directed east (at right angles to the axis of the loop). Note the current directions in (a): out of the page in M and into the page in L. With reference to the up–down, W–E cross arrows in (a):

(a)



(b)



8 When LM is aligned horizontally (L1–M1), what is the direction of the magnetic force on:
- **a** side L
- **b** side M?

Chapter review *continued*

9 In what direction, as seen from Y, will the loop rotate?

10 When LM is aligned vertically (L_2–M_2) what is the:
- a direction of the magnetic force on side L
- b direction of the magnetic force on side M
- c magnitude of the torque acting on the loop? Give a reason for your answer.

11 When LM is aligned vertically, which one of the following actions will result in a torque acting on the coil that will keep it rotating in an anticlockwise direction? (Assume it still has some momentum when it reaches the vertical position.)
- A decrease the current through the loop
- B increase the magnetic field strength
- C reverse the direction of the current through the coil

12 Briefly explain the function of the commutator in an electric motor.

13 Describe the basic set-up of cathode ray tubes and how the electrons are accelerated through the tubes.

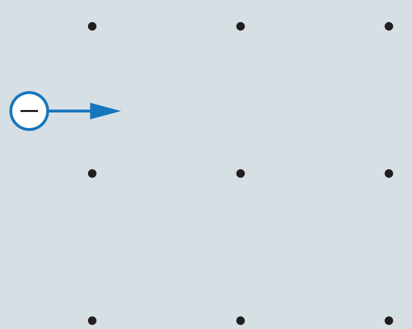
The following information relates to questions 14 and 15.

An electron-gun assembly emits electrons with energies of 10 keV. Ignore the effects of relativity when answering the following questions.

14 Calculate the magnitude of the predicted exit velocity of the electrons.

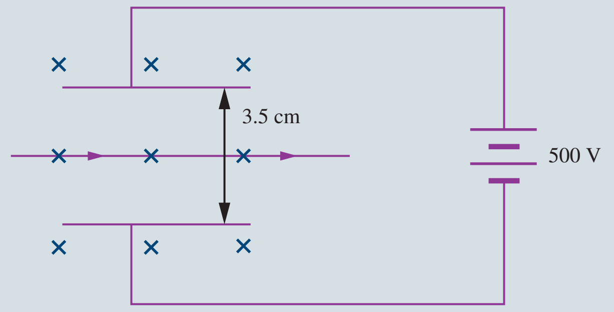
15 Upon exiting the electron-gun assembly, the electrons enter a uniform magnetic field of 1.5 T acting perpendicular to their motion. Calculate the predicted radius of the electron beam.

16 The diagram below represents an electron being fired at right angles towards a uniform magnetic field acting out of the page.



- a Copy the diagram and mark on it the continued path you expect the electron will follow.
- b Which factors would alter the path radius of the electron as it travels?

17 A stream of electrons travels in a straight line through a uniform magnetic field and between a pair of charged parallel plates, as shown in the diagram.



Calculate the:
- a electric field strength between the plates
- b speed of the electrons, given that the magnetic field is of strength 1.5×10^{-3} T.

18 Electrons in a cathode ray tube are accelerated through a potential difference from a cathode to a screen. Calculate the speed at which they hit the screen if the potential difference between electrodes is 4.5 kV.

19 An electron with speed of 4.3×10^6 m s^{-1} travels through a uniform magnetic field and follows a circular path of diameter 8.4×10^{-2} m. Calculate the magnetic field strength through which the electron travels.

20 a Calculate the force exerted on an electron travelling at speed of 6.4×10^6 m s^{-1} at right angles to a uniform magnetic field of strength 9.1×10^{-3} T.
- b Given that this force directs the electron in a circular path, calculate the radius of its orbit.

REVIEW QUESTIONS

How do things move without contact?

1 Draw the electric field pattern in the space between and around these two charges.

2 Two charges of +5 μC and −7 μC are positioned 0.4 m apart in air. What is the force that acts between them?

3 Calculate the electric field strength and direction at a distance of 3.5 mm directly to the left of a charge of +9.4 μC.

The following information relates to questions 4–7.

Two parallel plates have a distance of 3.8 cm between them and have a potential difference of 400 V across them.

4 What is the size of the electric field strength between the plates?

An electron is placed next to the negative plate.

5 What size force will be exerted on the electron?

6 What amount of energy in joules will be gained by the electron?

7 What is the speed of the electron by the time it reaches the positive plate?

8 Two metal plates have an electric field of 300 V m^{-1} between them and a separation distance of 12 cm. What is the voltage across the plates?

The following information relates to questions 9–11.

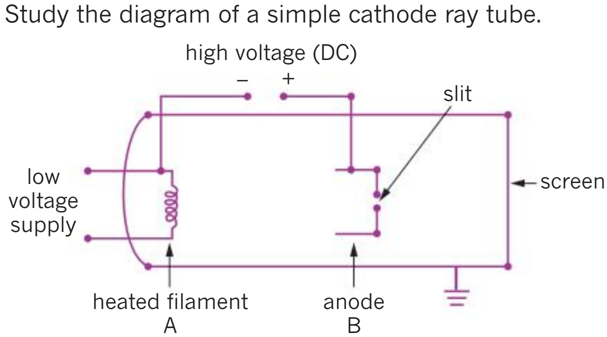
In a Millikan oil drop experiment, an oil drop of mass 1.96×10^{-14} kg is stationary between two horizontal parallel plates. The plates have a separation distance of 1.6 mm with 240 V between them.

9 Determine the size and direction of the electric field strength between the plates.

10 Calculate the size of the charge that must exist on the oil drop.

11 How many excess electrons are on the drop?

12 Study the diagram of a simple cathode ray tube.



What is the source of electrons in this device?

A the heated filament at A
B the positive anode at B
C the wires used in the circuit
D the screen used in the circuit

13 A particular electron gun accelerates an electron across a potential difference of 15 kV, a distance of 12 cm between a pair of charged plates. Calculate the magnitude of the force acting on the electron.

14 In an electron gun, an electron is accelerated by a potential difference of 28 kV. At what speed will the electrons exit the assembly?

15 If the electron in the previous question was accelerated a distance of 20 cm between a pair of charged parallel plates, then calculate the size of the electric field strength acting on the electron.

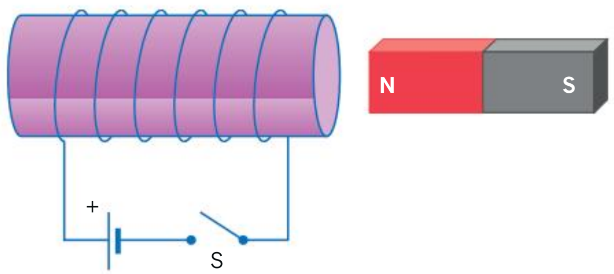
16 The left diagram below represents two conductors, both perpendicular to the page and both carrying equal currents into the page (shown by the crosses in the circles). In these questions ignore any contribution from the Earth's magnetic field. Choose the correct options from the arrows A–D and letters E–G.

E Out of page
F Into page
G Zero field

What is the direction of the magnetic field due to the two currents at each of the following points?

a point P
b point Q
c point R

17 An electromagnet with a soft iron core is set up as shown in the diagram below. A small bar magnet with its north end towards the electromagnet is placed to the right of it. The switch S is initially open. The following questions refer to the force between the electromagnet and the bar magnet under different conditions.



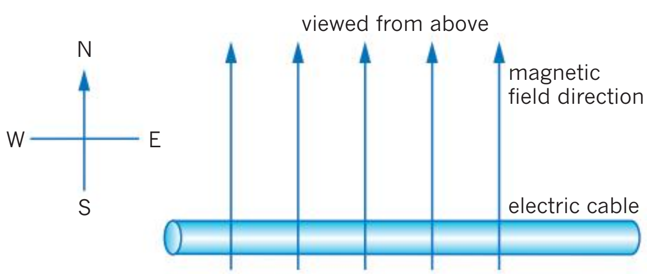
a Describe the force on the bar magnet while the switch remains open.

b Describe the force on the bar magnet when the switch is closed and a heavy current flows.

c The battery is removed and then replaced so that the current flows in the opposite direction. Describe the force on the bar magnet now when the switch is closed.

The following information applies to questions 18–22.

The diagram below shows a horizontal, east–west electric cable, located in a region where the magnetic field of the Earth is horizontal and has a magnitude of 1.0×10^{-5} T. The cable has a mass of 0.05 kg m^{-1}.



18 What is the magnitude of the magnetic force on a 1.0 m section of this cable if a 100 A current is flowing through it?

19 What is the direction of the current that will produce a force vertically upwards on this cable?

20 What magnitude of current would be required to produce zero resultant vertical force on a 1.0 m section of this cable?

21 Assume that a 100 A current is flowing through this cable from west to east. What would be the magnitude of the change in magnetic force per metre on this cable if the direction of this current was reversed?

22 The cable is no longer at right angles, but makes an angle θ with the direction of the Earth's magnetic field. What force would 100 A current passing through this cable produce?

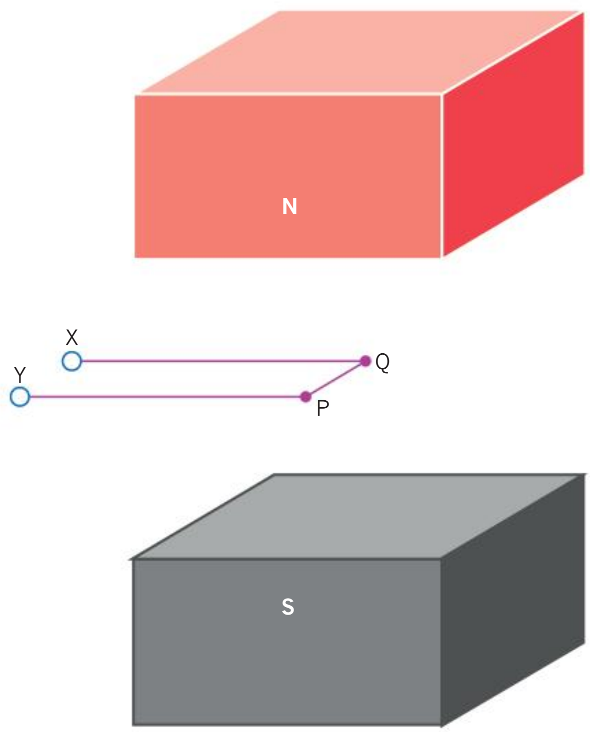
A the same magnetic force on the cable as when it was horizontal

B a smaller magnetic force than when it was horizontal

C a larger magnetic force than when it was horizontal

The following information applies to questions 23–26.

The following diagram shows a section of a conducting loop XQPY, part of which is placed between the poles of a magnet whose uniform field strength is 1.0 T. The side PQ has length 5.0 cm. X is connected to the positive terminal of a battery while Y is connected to the negative terminal. A current of 1.0 A then flows through this loop.



23 What is the magnitude of the force on side PQ?

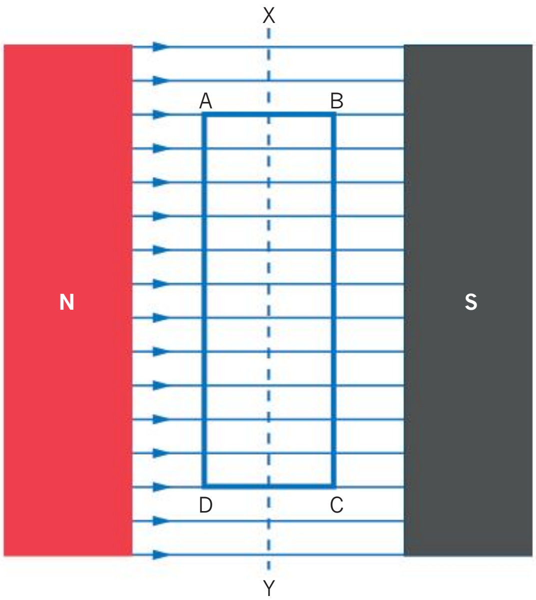
24 What is the direction of the force on side PQ?

25 What is the magnitude of the force on a 1.0 cm section of side XQ that is located in the magnetic field?

26 The direction of the current through the loop is reversed by connecting X to the negative terminal and Y to the positive terminal of the battery. What is the direction of the force on side PQ?

The following information relates to questions 27–29.

A rectangular coil containing 100 turns and dimensions 10 cm × 5 cm is located in a magnetic field $B = 0.25$ T, as shown. It is free to rotate about the axis XY. The coil carries a constant current $I = 200$ mA flowing in the direction ADCB.



27 What is the magnitude and direction of the magnetic force on the following sides?

a AB

b DC

28 What is the magnitude and direction of the magnetic force on the following sides?

A AD

B BC

29 Describe the likely motion of the coil if it is free to rotate.

30 A student constructs a simple DC electric motor consisting of *N* loops of wire wound around a wooden armature, and a permanent horseshoe-shaped magnetic of strength *B*. The student connects the motor to a 9 V battery but is not happy with the speed of rotation of the armature. Which one or more of the following modifications will most likely increase the speed of rotation of the armature?

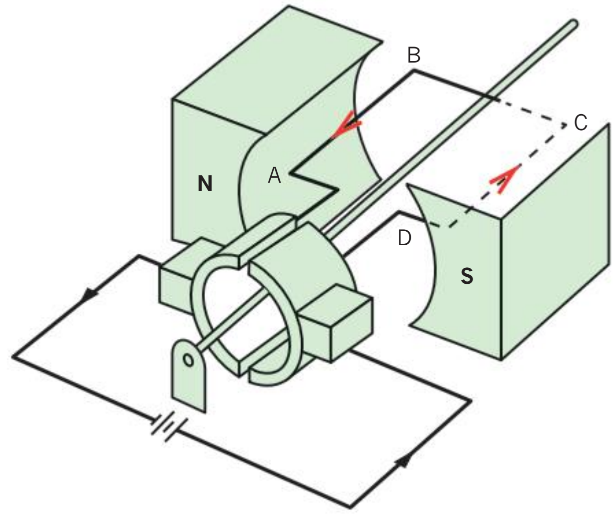
A increase the number of turns *N*

B use a 12 V battery instead of a 9 V battery

C replace the wooden armature with one of soft iron

D connect a 100 Ω resistor in series with the armature windings

31 Consider the electric motor shown.



a The direction of the current in the coil is shown (from D anticlockwise to A). What is the direction of the force on sides AB and CD?

b In what position of the coil is the turning effect of the forces greatest?

c At one point in the rotation of the coil the turning effect becomes zero. Explain where this occurs and why the motor actually continues to rotate.

32 A rectangular loop of 100 turns is suspended in a magnetic field $B = 0.50$ T. The plane of the loop is parallel to the direction of the field. The dimensions of the loop are 20 cm perpendicular to the field lines and 10 cm parallel to them.

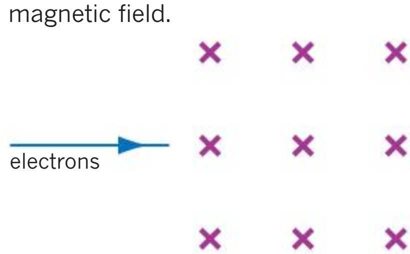
a It is found that there is a force of 40 N on each of the sides perpendicular to the field. What is the current in each turn of the loop?

b This loop is then replaced by a square loop of 10 cm each side, with twice the current and half the number of turns. What is the force on each of the perpendicular sides now?

c The original rectangular loop with the original current is returned but a new magnet is found which provides a field strength of 0.80 T. What is the force on the 20 cm side now?

33 If an electron travels through a magnetic field of strength 1.2 T with a speed of 4.2×10^6 m s^{-1}, calculate the radius of the path it will follow.

34 This diagram shows a stream of electrons entering a magnetic field. Reproduce the diagram and show the subsequent path of the electrons through the magnetic field.



35 An electron beam travelling through a cathode ray tube is subjected to simultaneous electric and magnetic fields. The electrons emerge with no deflection. Given that the potential difference across the parallel plates X and Y is 3 kV, and that the applied magnetic field is of strength 1.6×10^{-3} T, calculate the distance between plates X and Y.

36 Use Newton's law of universal gravitation to calculate the size of the force between two masses of 24 kg and 81 kg, with a distance of 0.72 m between their centres.

The following information relates to questions 37–40.

Consider an astronaut inside a spacecraft from launch to a stable orbit. Choose your answers to questions 37 to 40 from the following options:

- **A** apparent weightlessness
- **B** weightlessness
- **C** apparent weight
- **D** gravitational force
- **E** none of the above

37 As the astronaut and spacecraft are launched which of the above will be greater than normal?

38 As the astronaut and spacecraft are launched, which of the above will remain constant?

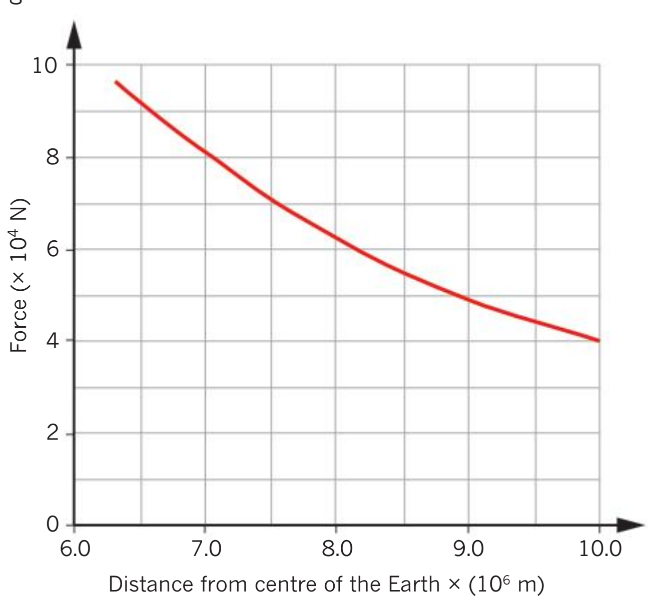
39 As the astronaut and spacecraft are in a stable orbit above the Earth, which of the above will apply to the astronaut?

40 If the astronaut and spacecraft ventured into deep space, which of the above would apply to the astronaut?

The following information applies to questions 41–44.

A 10 000 kg spacecraft is drifting directly towards the Earth. When it is at an altitude of 600 km, its speed is 1.5 km s^{-1}. The radius of the Earth is 6400 km.

The following graph shows the force on the spacecraft against distance from the Earth.



41 How much gravitational potential energy would the spacecraft lose as it falls to a distance of 6500 km?

42 Determine the speed of the spacecraft at a distance of 6500 km.

43 What is the weight of the spacecraft when it is at an altitude of:

a 3600 km?

b 6.0×10^5 m?

44 How does the acceleration of the spacecraft change as it moves from an altitude of 600 km to an altitude of 100 km? Include numerical data in your answer.

The following information applies to questions 45–47.

The ATV2 satellite was launched by the European Space Agency in February 2011 to deliver supplies to the International Space Station. The ATV2 satellite is in a circular orbit of radius 6.73×10^6 m.

The following information may be required to answer these questions.

Mass of ATV2 satellite and cargo = 1.2×10^4 kg

Mass of Earth = 5.98×10^{24} kg

Radius of Earth = 6.37×10^6 m

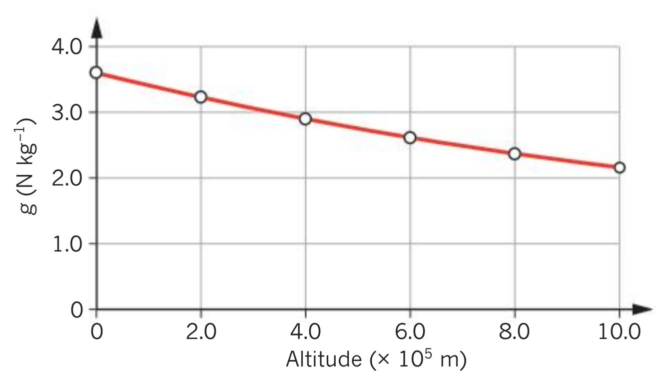
45 What is the weight of the ATV2 and cargo when it is in its orbit?

46 Calculate the orbital period of the ATV2 satellite.

47 The satellite delivers its cargo to the ISS and now orbits with a mass of just 6.0 tonnes. How does this reduced mass affect the orbital period of the ATV2?

The following information applies to questions 48–50.

A small asteroid has just smashed into the surface of Mars and a lump of Martian rock of mass 20 kg has been thrown into space with 40 MJ of kinetic energy. A graph of gravitational field–distance from the surface of Mars is shown below.



48 What is the gravitational force acting on the Martian rock when it is at an altitude of 300 km?

49 How much kinetic energy (in MJ) does the rock lose as it travels from the surface of Mars to an altitude of 6.0×10^5 m?

50 The rock eventually comes to a stop and starts to fall back towards Mars. Without actually doing the calculations, explain how you would determine the altitude at which the rock stopped.

CHAPTER 04

Electromagnetic induction and transmission of electricity

In this chapter, electromagnetic induction—the creation of an electric current from a changing magnetic flux—is explored.

In 1831, Englishman Michael Faraday and American Joseph Henry independently discovered that a changing magnetic flux could induce an electric current in a conductor. This discovery made possible the production of vast quantities of electricity. Today, whether the primary energy source is burning coal, wind, nuclear fission or falling water, the vast bulk of the world's electrical energy production is the result of electromagnetic induction.

Key knowledge

At the end of this chapter you will have investigated the generation of electricity by electromagnetic induction and the transmission of electricity. You will be able to:

- calculate magnetic flux when the magnetic field is perpendicular to the area, and describe the qualitative effect of differing angles between the area and the field; i.e. $\Phi_B = B_{\perp}A$
- investigate and analyse theoretically and practically the generation of electromotive force (emf) including AC voltage and calculations using induced emf: $\varepsilon = -N\frac{\Delta\Phi_B}{\Delta t}$ with reference to:
  - rate of change of magnetic flux
  - number of loops through which the flux passes
  - direction of induced emf in a coil
- explain the production of DC voltage in DC generators and AC voltage in alternators, including the use of split-ring commutators and slip rings respectively
- compare sinusoidal AC voltages produced as a result of the uniform rotation of a loop in a constant magnetic field with reference to frequency, period, amplitude, peak-to-peak voltage ($V_{p\text{–}p}$) and peak-to-peak current ($I_{p\text{–}p}$)
- compare alternating voltage expressed as the root-mean-square (rms) to a constant DC voltage developing the same power in a resistive component
- convert between rms, peak and peak-to-peak values of voltage and current
- analyse transformer action with reference to electromagnetic induction for an ideal transformer: $\frac{N_1}{N_2} = \frac{V_1}{V_2} = \frac{I_2}{I_1}$
- analyse the supply of power by considering transmission losses across transmission lines
- identify the advantage of the use of AC power as a domestic power supply.

4.1 Inducing an emf in a magnetic field

After Oersted's discovery that an electric current produces a magnetic field (see Section 2.3 on p45), Michael Faraday, an English scientist, was convinced that the reverse should also be true—a magnetic field should be able to produce an electric current.

Faraday wound two coils of wire onto an iron ring (see Figure 4.1.1). By connecting a battery to one of the coils he created a strong current through one coil which therefore created a strong magnetic field. He expected to then detect the creation of an electric current in the second coil. No matter how strong the magnetic field, he could not detect an electric current in the other coil.

One day he noticed that the galvanometer (a type of sensitive ammeter) attached to the second coil flickered when he turned on the current that created the magnetic field. It gave another flicker, in the opposite direction, when he turned the current off. It was not the strength of the magnetic field that mattered, but the change in the magnetic field.

The creation of an electric current in a conductor due to a change in the magnetic field acting on that conductor is now called **electromagnetic induction**. This section focuses on this concept.

FIGURE 4.1.1 Michael Faraday's original induction coil. Passing a current through one coil induces a voltage in the second coil by a process called mutual inductance. This coil is now on display at the Royal Institution in London.

CREATING AN ELECTRIC CURRENT

In his attempts to produce an electric current from a magnetic field, Faraday had no success with a constant magnetic field but was able to observe the creation of an electric current whenever there was a change in the magnetic field. This current is produced by an induced emf, ε. Although the term emf is derived from the name electromotive force, it is a voltage, or potential difference, rather than a force. Figure 4.1.2 indicates the induction of emf, and therefore current, caused by the perpendicular movement of a conducting wire relative to a magnetic field.

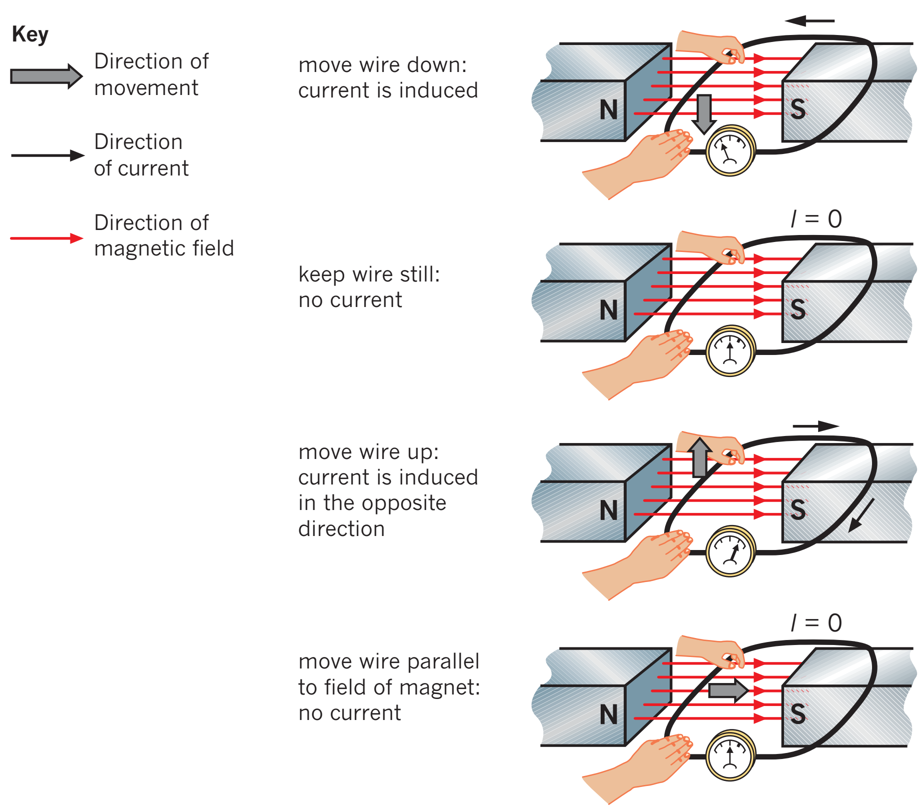


FIGURE 4.1.2 Electromotive force (emf) is induced in a wire when it moves perpendicular to a magnetic field.

PHYSICSFILE

Models and theories

Michael Faraday was not alone in the discovery of electromagnetic induction. Joseph Henry (1797–1878), an American physicist, independently discovered the phenomenon of electromagnetic induction a little ahead of Michael Faraday, but Faraday was the first to publish his results. Henry later improved the design of the electromagnet, using a soft iron core wrapped in many turns of wire. He also designed the first reciprocating electric motor. Henry is credited with first discovering the phenomenon of self-induction, and the unit of inductance is named after him. He also introduced the electric relay, which made the sending of telegrams possible. Henry was the first director of the Smithsonian Institution.

While Faraday will be largely referred to throughout this text, it is worth noting that there can be a number of contributors who together build on the understanding of key ideas. Joseph Henry's contributions should not be forgotten.

MAGNETIC FLUX

To be able to develop ideas about the change in a magnetic field that induces an emf which can then create (or induce) a current, it is useful to be able to describe the 'amount of magnetic field'. This amount of magnetic field is referred to as the **magnetic flux**, a scalar quantity, denoted by the symbol Φ_B. Faraday pictured a magnetic field as consisting of many lines of force. The density of the lines represents the strength of the magnetic field. Magnetic flux can be related to the total number of these lines that pass within a particular area. A strong magnetic field acting over a small area can produce the same amount of magnetic flux as a weaker field acting over a larger area, as shown in Figure 4.1.3. For this reason, magnetic field strength, B, is also referred to as **magnetic flux density**. B can be thought of as being proportional to the number of magnetic field lines per unit area perpendicular to the magnetic field. The magnetic flux will be at a maximum when the area examined is perpendicular to the magnetic field and zero when the area being examined is parallel to the magnetic field.

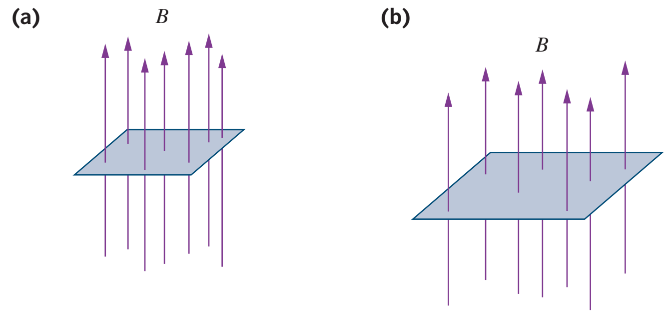


FIGURE 4.1.3 Magnetic flux. A strong magnetic field acting over a small area (a) will have the same magnetic flux as a weaker magnetic field acting over a larger area (b).

Based on this, magnetic flux is defined as the product of the strength of the magnetic field, B, and the area of the field perpendicular to the field lines, i.e.

$$\Phi_B = B_{\perp}A$$

where Φ_B, or simply Φ, is the magnetic flux. The unit for magnetic flux is the weber, abbreviated to Wb, where 1 Wb = 1 T m^{-2}

B is the strength of the magnetic field in tesla (T)

A is the area perpendicular to the magnetic field, measured in square metres (m^2).

The subscript $_{\perp}$ is included in the formula to indicate that the area referred to is perpendicular to the magnetic field.

Since it is the area perpendicular to the magnetic field, the angle between the magnetic field and the area through which the field passes will obviously affect the amount of magnetic flux. As the angle increases or decreases from 90° the amount of magnetic flux will decrease, until reaching zero when the area under consideration is parallel to the magnetic field. Referring to Figure 4.1.4, then the relationship between the amount of magnetic flux and the angle θ to the field is:

$$\Phi_B = BA\cos\theta$$

It is important to note that θ is not the angle between the plane of the area and the magnetic field. Rather, it is the angle between a normal to the area and the direction of the magnetic field; hence the use of cos θ. When the area is at right angles to the magnetic field, the angle θ between the normal and the field is 0° and cos 0° = 1 (top diagram in Figure 4.1.4). When the area is parallel to the magnetic field, the angle θ between the normal and the field is 90° and cos 90° = 0 (bottom diagram in Figure 4.1.4).

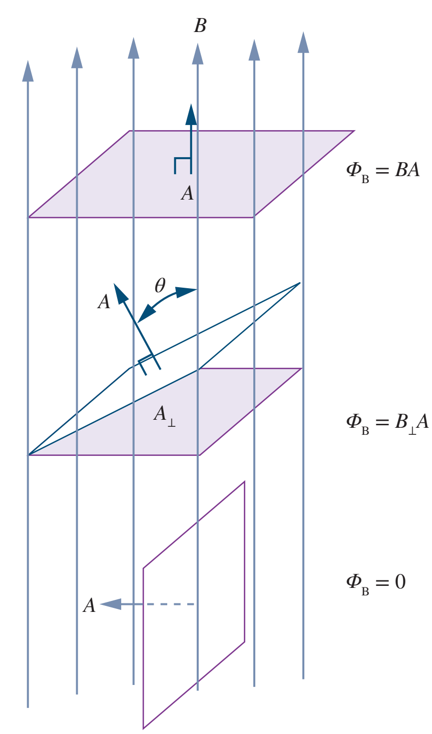


FIGURE 4.1.4 The magnetic flux is the strength of the magnetic field, B, multiplied by the area perpendicular to the magnetic field, given by $A\cos\theta$ and shown as the shaded areas in the above diagrams.

Worked example 4.1.1

MAGNETIC FLUX

A student places a horizontal square coil of wire of side length 5.0 cm into a uniform vertical magnetic field of 0.10 T. How much magnetic flux 'threads' the coil?	
Thinking	**Working**
Calculate the area of the coil perpendicular to the magnetic field.	side length = 5.0 cm = 0.05 m area of the square = $(0.05\ \text{m})^2$ = $0.0025\ \text{m}^2$
Calculate the magnetic flux.	$\Phi_B = B_{\perp}A$ $= 0.1 \times 0.0025$ $= 0.00025$ Wb
State the answer in an appropriate form.	$\Phi_B = 2.5 \times 10^{-4}$ Wb or 0.25 mWb

Worked example: Try yourself 4.1.1

MAGNETIC FLUX

A student places a horizontal square coil of wire of side length 4.0 cm into a uniform vertical magnetic field of 0.050 T. How much magnetic flux 'threads' the coil?

Note that in Worked example 4.1.1 an area of 5 cm × 5 cm = 25 cm^2 was considered, and this corresponds to 0.0025 m^2 or 25×10^{-4} m^2, so:

$$1\ \text{cm}^2 = 1 \times 10^{-4}\ \text{m}^2$$

THE INDUCED EMF IN A MOVING CONDUCTOR

It was discovered that a change in the magnetic field, when a magnet is moved closer to a conductor, leads to an induced emf that in turn produces an **induced current**. While Faraday largely based his investigations on induced currents in coils, another way of inducing an emf is by moving a straight conductor in a magnetic field. It's not hard to understand why this is the case, when you know that charges moving in a magnetic field will experience a force.

In Chapter 2, it was established that when a charge, q, moves at a speed, v, perpendicular to a magnetic field, B, the charge experiences a force, F, equal to qvB. So:

$$F = qvB$$

Considering the direction of movement shown in Figure 4.1.5, the force on the positive charges within the moving conductor would be along the conductor and out of the page. The force on the negative charges within the conductor would be along the conductor but into the page.

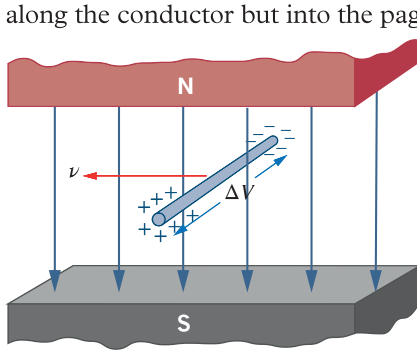


FIGURE 4.1.5 A potential difference, ΔV, will be produced across a straight wire moving to the left in a downward-pointing magnetic field.

As the charges in Figure 4.1.5 move apart due to the force they are experiencing from the magnetic field, one end of the conductor will become more positive, the other more negative and a potential difference, ΔV, or emf will be induced between the ends of the conductor.

Consider now an electron moving along the conductor. The force from the magnetic field will do work on the electron as it moves along a length, l. To calculate the work done:

W = force × distance = $qvB \times l$

The emf is equal to the work done per unit charge, so:

$$\varepsilon = \frac{W}{q} = \frac{qlvB}{q}$$

and thus:

$\varepsilon = lvB$

where ε is the induced emf (V)

l is the length of the conductor (m)

v is the speed of the conductor perpendicular to the magnetic field (m s^{-1})

B is the strength of the magnetic field (T)

Worked example 4.1.2

ELECTROMOTIVE FORCE ACROSS AN AIRCRAFT'S WINGS

Will a moving airplane develop a dangerous emf between its wing tips solely from the Earth's magnetic field? An aircraft with a wing span of 64 m is flying at a speed of 990 km h^{-1} at right angles to the Earth's magnetic field of 5.0×10^{-5} T.

Thinking	Working
Identify the quantities required in the correct units.	$\varepsilon = ?$ $l = 64$ m $B = 5.0 \times 10^{-5}$ T $v = 990$ km h^{-1} $= 990 \times \frac{1000}{3600}$ $= 275$ m s^{-1}
Substitute into the appropriate formula and simplify.	$\varepsilon = lvB$ $= 64 \times 275 \times 5.0 \times 10^{-5}$ $= 0.88$ V
State your answer as a response to the question.	$\varepsilon = 0.88$ V This is a very small emf and would not be dangerous.

Worked example: Try yourself 4.1.2

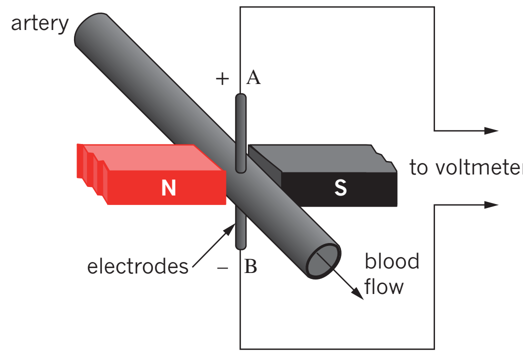
ELECTROMOTIVE FORCE ACROSS AN AIRCRAFT'S WINGS

Will a moving airplane develop a dangerous emf between its wing tips solely from the Earth's magnetic field? A fighter jet with a wing span of 25 m is flying at a speed of 2000 km h^{-1} at right angles to the Earth's magnetic field of 5.0×10^{-5} T.

Worked example 4.1.3

FLUID FLOW MEASUREMENT

The rate of fluid flow within a vessel can be measured using the induced emf when the fluid contains charged ions. A small magnet and sensitive voltmeter calibrated to measure speed are used. This can be applied to measure the flow of blood (which contains iron in solution) in the human body. If the diameter of a particular artery is 2.00 mm, the strength of the magnetic field applied is 0.10 T and the measured emf is 0.10 mV, what is the speed of the flow of the blood within the artery?

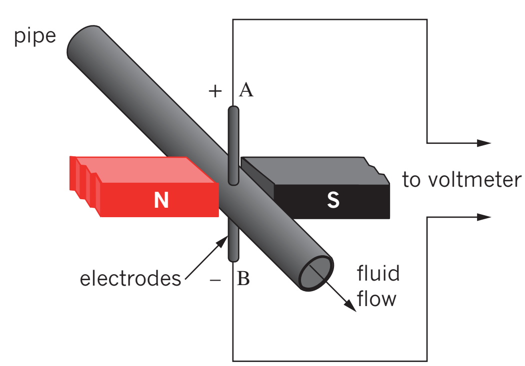


Thinking	Working
Identify the quantities required and put them into SI units.	$\varepsilon = 0.10$ mV $= 1.0 \times 10^{-4}$ V $l = 2.00$ mm $= 2.00 \times 10^{-3}$ m $v = ?$ $B = 0.10$ T
Rearrange the appropriate formula, substitute and simplify.	$\varepsilon = lvB$ $v = \frac{\varepsilon}{lB}$ $= \frac{1.0 \times 10^{-4}}{2.00 \times 10^{-3} \times 0.10}$ $= 0.50$ m s^{-1}
State your answer with the correct units.	$v = 0.50$ m s^{-1}

Worked example: Try yourself 4.1.3

FLUID FLOW MEASUREMENT

The rate of fluid flow within a vessel can be measured using the induced emf when the fluid contains charged ions. A small magnet and sensitive voltmeter calibrated to measure speed are used. This can be applied to measure fluid flow in small pipes. If the diameter of a particular small pipe is 1.00 cm, the strength of the magnetic field applied is 0.10 T and the measured emf is 0.50 mV, what is the speed of the fluid flow?



4.1 Review

SUMMARY

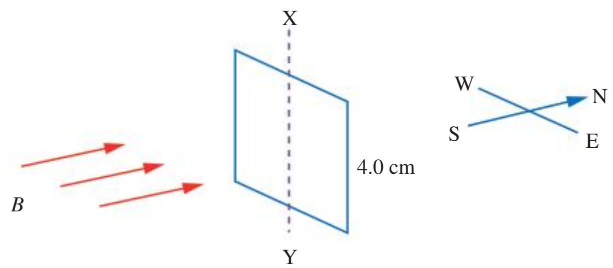
- An induced emf, ε, is produced by a changing magnetic flux in a process called electromagnetic induction.
- Magnetic flux is defined as the product of the strength of the magnetic field, B, and the area of the field perpendicular to the field lines, i.e. $\Phi_B = B_{\perp}A$.
- The amount of magnetic flux varies with the angle of the field to the area under investigation. It is a maximum when the area is perpendicular (90°) and zero when the area is parallel to the field. i.e. $\Phi_B = BA \cos \theta$.
- The unit for magnetic flux is the weber, Wb; 1 Wb = 1 T m^{-2}.
- The induced emf in a straight conductor moving in a magnetic field, B, is given by $\varepsilon = lvB$, where $v \perp B$.

KEY QUESTIONS

1 Which of the following scenarios will *not* induce an emf in a long, straight conductor?
   A A magnet is stationary alongside the conductor.
   B A magnet is brought near the conductor.
   C The conductor is brought into a magnetic field.
   D The conductor is rotated within a magnetic field.

2 A student places a 4.0 cm square coil of wire parallel to a uniform vertical magnetic field of 0.050 T. How much magnetic flux 'threads' the coil?

3 A square loop of wire, of side 4.0 cm, is in a region of uniform magnetic field, $B = 2.0 \times 10^{-3}$ T north, as in the diagram below. The loop is free to rotate about a vertical axis XY. When the loop is in its initial position, its plane is perpendicular to the direction of the magnetic field. What is the magnetic flux passing through the loop?



4 The same square loop of wire as in Question 3 is initially perpendicular to the magnetic field. The loop is free to rotate about a vertical axis XY.
Describe what happens to the amount of magnetic flux passing through the loop as the loop is rotated through one complete revolution.

5 A circular coil of wire, of radius 5.0 cm, is perpendicular to a region of uniform magnetic field, B = 1.6 mT. What is the magnetic flux passing through the loop?

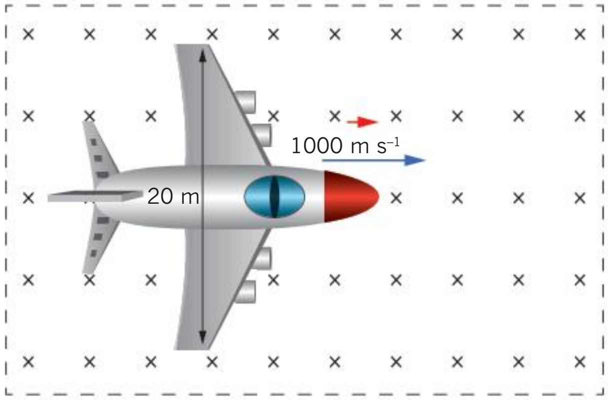
6 A moving rod 12 cm long is being moved at a speed of 0.150 m s^{-1} perpendicular to a magnetic field, B. If the strength of the magnetic field is 0.800 T, calculate the induced emf in the rod.

7 A metal rod is 13.2 cm long. It generates an emf of 100 mV while moving perpendicular to a magnetic field of strength 0.90 T. At what speed is it moving?

8 A metal rod generates an emf of 80 mV while moving at a speed of 1.6 m s^{-1} perpendicular to a magnetic field of strength 0.50 T. How long is the metal rod?

9 A rod of length 10 cm and very small diameter is held vertically and dropped downwards through a vertically-upwards-directed magnetic field of strength 0.80 T. If the rod's initial speed was zero, what would be the induced emf in the rod at an instant 5.0 s after it was dropped?

10 Calculate the magnitude of the induced emf between the ends of the wings of an aircraft whose wingspan is 20 m, given that the aircraft is moving at a speed of 1000 m s^{-1} in the magnetic field of the Earth in a plane perpendicular to the lines of the field, where the flux density is 2.5×10^{-5} T.



4.2 Induced emf from a changing magnetic flux

Faraday's early experiments largely centred on investigating electromagnetic induction in coils, or multiple loops, of wire. Faraday found that if a magnet is quickly moved into a coil, an emf is induced which causes a current to flow in the coil. If the magnet is removed, then a current flows in the coil in the opposite direction. Alternatively, if the magnet is held steady and the coil is moved in such a way that changes the magnetic flux, then once again an emf is induced and an electric current flows. It doesn't matter whether the coil or the magnet is moved—it is a *change* in flux that is required to induce the emf (see Figure 4.2.1). This discovery led Faraday to his law of induction. Faraday's law of induction is the focus of this section.

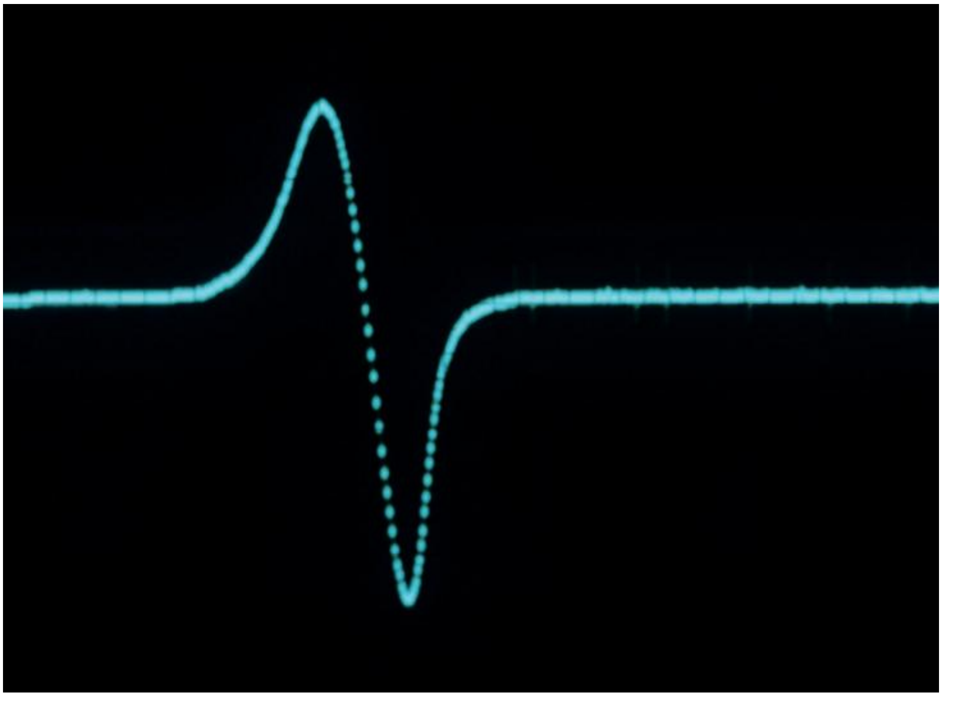


FIGURE 4.2.1 Oscilloscope trace from an electric coil, showing the voltage across the coil as a magnet is dropped through it.

FACTORS AFFECTING INDUCED EMF

Faraday quantitatively investigated the factors affecting the size of the emf induced in a coil. Firstly, emf will be induced by a change in the magnetic field. A simple example of this is to witness the emf induced when a magnet is brought towards or away from a wire coil. The greater the change, the greater the emf.

However, it is not only a change in the strength of a magnetic field, B, that induces an emf. It was noticed that an emf can be induced by changing A, the area perpendicular to the magnetic field through which the magnetic field lines pass, while keeping B constant. An example of this is to witness the emf induced when a wire coil is rotated in the presence of a fixed magnetic field. This discovery indicates that the requirement for an induced emf is to have a *changing magnetic flux*, Φ_B.

Finally, Faraday discovered that the faster the change in magnetic flux, the greater the induced emf. This can be seen in the oscilloscope trace of a magnet falling through a coil as shown in the Figure 4.2.1. The magnet is accelerated by gravity as it drops through the coil. Hence, the peak emf induced when the magnet first enters the coil at a relatively lower speed is noticeably less than the peak emf when the magnet leaves the coil at a faster speed. Thus, it is the *rate of change* of magnetic flux that determines the induced emf.

FARADAY'S LAW OF INDUCTION

Faraday's investigations led him to conclude that the average emf induced in a conducting loop, in which there is a changing magnetic flux, is proportional to the rate of change of flux.

This is now known as **Faraday's law** of induction and is one of the basic laws of electromagnetism.

Magnetic flux is defined as $\Phi_B = B_\perp A$.

If the flux through N turns (or loops) of a coil changes from Φ_1 to Φ_2 during a time t, then the average induced emf during this time will be:

$$\varepsilon = -N\frac{(\Phi_2 - \Phi_1)}{t}$$

and if the change in magnetic flux $\Phi_2 - \Phi_1 = \Delta\Phi_B$, then

$$\varepsilon = -N\frac{\Delta\Phi_B}{\Delta t}$$

The negative sign is placed there as a reminder of the direction of the induced emf. This is discussed further in the next section 'Lenz's law and its applications'. Normally you will be concerned only with the magnitude of the emf, which means you don't consider the negative sign or any negative quantities in a calculation.

If the ends of the coil are connected to an external circuit, then a current, I, will flow. The magnitude of the current is found using Ohm's law, which is:

$$I = \frac{V}{R}$$

where R is the resistance and ΔV is the emf of the coil.

A coil not connected to a circuit will act like a battery not connected to a circuit. There will still be an induced emf but no current will flow.

Worked example 4.2.1

INDUCED EMF IN A COIL

A student winds a coil of area 40 cm^2 with 20 turns. He places it horizontally in a vertical uniform magnetic field of 0.10 T.

a Calculate the magnetic flux perpendicular to the coil.	
Thinking	**Working**
Identify the quantities to calculate the magnetic flux through the coil and convert to SI units where required.	$\Phi_B = B_\perp A$ $B = 0.10$ T $A = 40$ cm^2 $= 40 \times 10^{-4}$ m^2
Calculate the magnetic flux and give your answer with appropriate units.	$\Phi_B = B_\perp A = 0.10 \times 40 \times 10^{-4}$ $= 4.0 \times 10^{-4}$ Wb

b Calculate the magnitude of the average induced emf in the coil when the coil is removed from the magnetic field in a time of 0.5 s.	
Identify the quantities for determining the induced emf. Ignore the negative sign.	$\varepsilon = -N\frac{\Delta\Phi_B}{\Delta t}$ $N = 20$ turns $\Delta\Phi_B = \Phi_2 - \Phi_1$ $= 0 - 4.0 \times 10^{-4}$ $= 4.0 \times 10^{-4}$ Wb $\Delta t = 0.5$ s
Calculate the magnitude of the average induced emf, ignoring the negative sign that indicates the direction.	$\varepsilon = -N\frac{\Delta\Phi_B}{\Delta t}$ $= 20 \times \frac{4.0 \times 10^{-4}}{0.5}$ $= 0.016$ V

PHYSICSFILE

Microphones

Many microphones operate by taking advantage of Faraday's law of induction. The so-called 'dynamic' microphone uses a tiny coil attached to a diaphragm, which vibrates with the sound. The coil vibrates within the magnetic field of a permanent magnet, thus producing an induced emf that varies with the original sound.

Worked example: Try yourself 4.2.1

INDUCED EMF IN A COIL

A student winds a coil of area 50 cm^2 with 10 turns. She places it horizontally in a vertical uniform magnetic field of 0.10 T.

a Calculate the magnetic flux perpendicular to the coil.

b Calculate the magnitude of the average induced emf in the coil when the coil is removed from the magnetic field in a time of 1.0 s.

Worked example 4.2.2

NUMBER OF TURNS IN A COIL

A coil of cross-sectional area 1.0×10^{-3} m^2 experiences a change in the strength of a magnetic field from 0 to 0.20 T over 0.50 s. If the magnitude of the average induced emf is measured as 0.10 V, how many turns must be on the coil?

Thinking	Working
Identify the quantities needed to calculate the magnetic flux through the coil when in the presence of the magnetic field and convert to SI units where required.	$\Phi_B = B_{\perp}A$ $B = 0.20$ T $A = 1.0 \times 10^{-3}$ m^2
Calculate the magnetic flux when in the presence of the magnetic field.	$\Phi_B = B_{\perp}A$ $= 0.20 \times 1.0 \times 10^{-3}$ $= 2.0 \times 10^{-4}$ Wb
Identify the quantities from the question required to complete Faraday's law.	$\varepsilon = -N\frac{\Delta\Phi_B}{\Delta t}$ $N = ?$ $\Delta\Phi_B = \Phi_2 - \Phi_1$ $= 2.0 \times 10^{-4} - 0$ $= 2.0 \times 10^{-4}$ Wb $\Delta t = 0.50$ s $\varepsilon = 0.10$ V
Rearrange Faraday's law and solve for the number of turns on the coil. Ignore the negative sign.	$\varepsilon = -N\frac{\Delta\Phi_B}{\Delta t}$ $N = -\frac{\varepsilon\Delta t}{\Delta\Phi_B}$ $= \frac{0.10 \times 0.50}{2.0 \times 10^{-4}}$ $= 250$ turns

Worked example: Try yourself 4.2.2

NUMBER OF TURNS IN A COIL

A coil of cross-sectional area 2.0×10^{-3} m^2 experiences a change in the strength of a magnetic field from 0 to 0.20 T over 1.00 s. If the magnitude of the average induced emf is measured as 0.40 V, how many turns must be on the coil?

4.2 Review

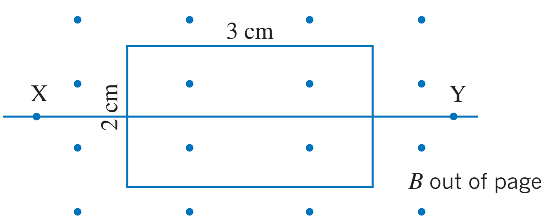
SUMMARY

- The emf induced in a conducting loop in which there is a changing magnetic flux is proportional to the negative rate of change of flux.
- $\varepsilon = -N\frac{\Delta\Phi_B}{\Delta t}$, which is Faraday's law of induction.
- The negative sign in Faraday's law indicates direction. For questions involving only magnitudes, you should not use the negative sign or any negative quantities.

KEY QUESTIONS

The following information relates to questions 1–3.

A single rectangular wire loop is located with its plane perpendicular to a uniform magnetic field of 2.0 mT, directed out of the page, as shown below. The loop is free to rotate about a horizontal axis XY.

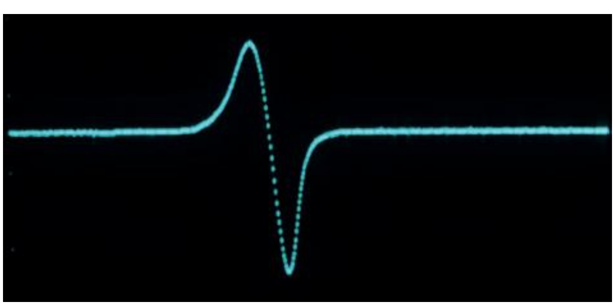


1 How much magnetic flux is threading the loop in this position?

2 The loop is rotated about the axis XY, through an angle of 90°, so that its plane becomes parallel to the magnetic field. How much flux is threading the loop in this new position?

3 If the loop completes one-quarter of a rotation in 40 ms, what is the average induced emf in the loop?

4 When a magnet is dropped through a coil, a voltage sensor will detect an induced voltage in the coil as shown below.



The area under the curve above zero is exactly equal to the area above the curve below zero because:

A The strength of the magnet is the same.

B The area of the coil is the same.

C The strength of the magnet and area of the coil are the same.

D The magnet speeds up as it falls through the coil.

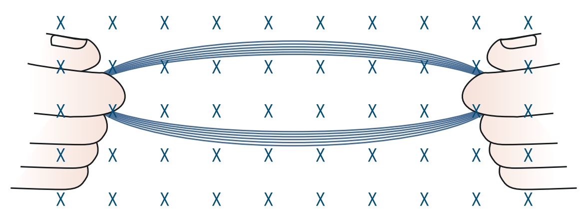
The following information relates to questions 5 and 6.

A coil of 500 turns, each of area 10 cm^2, is wound around a square frame. The plane of the coil is initially parallel to a uniform magnetic field of 80 mT. The coil is then rotated through an angle of 90° so that its plane becomes perpendicular to the field. The rotation is completed in 20 ms.

5 What is the average emf induced in each turn during this time?

6 What is the average emf induced in the coil in Question 5 during the time the coil rotated?

7 A student stretches a flexible wire coil of 30 turns and places it in a uniform magnetic field of strength 5.0 mT, directed into the page, as shown. While it is in the field, the student allows the coil to regain its original shape. In doing so, the area of the coil changes at a constant rate from 250 cm^2 to 50 cm^2 in 0.50 s.



Find the average emf induced in the coil during this time.

8 A student has a flexible wire coil of variable area of 100 turns and a strong bar magnet, which has been measured to produce a magnetic field of strength B = 100 mT, a short distance from it. She has been instructed to demonstrate electromagnetic induction by using this equipment to light up an LED rated at 1.0 V. Explain, including appropriate calculations, one method with which she could complete this task.

9 A wire coil consisting of a single turn is placed perpendicular to a magnetic field that experiences a decrease in strength of 0.10 T in 0.050 s. If the emf induced in the coil is 0.020 V, what is the area of the coil?

10 A wire coil consisting of 100 turns with an area of 50 cm^2 is placed inside a vertical magnetic field of strength 0.40 T, and then rotated about a horizontal axis. For each quarter turn, the average emf induced in the coil is 1600 mV. Calculate the time taken for a quarter turn of the coil.

4.3 Lenz's law and its applications

In the previous section Faraday's law of induction was introduced, including the equation:

$$\varepsilon = -N\frac{\Delta\Phi_B}{\Delta t}$$

The negative sign is there to remind you in which direction the induced emf acts—that is, in which direction current flows as a result of the induced emf.

Lenz's law, which is the focus of this section, is a common way of understanding how electromagnetic induction obeys the principles of conservation of energy and explains the direction of the induced emf. It is named after Heinrich Lenz, whose research put a definite direction to the current created by the induced emf resulting from a changing magnetic flux.

Understanding the direction of the current resulting from an induced emf and how it is produced has allowed electromagnetic induction to be used in a vast array of devices that have transformed modern society, in particular in electrical generators. A metal detector is another example of a device that uses Lenz's law (see Figure 4.3.1).

FIGURE 4.3.1 A diver using a metal detector. If a metal object is found underneath the coil of the detector, an emf will be induced which creates a current that will affect the original current. The direction of the induced current is predicted by using Lenz's law.

THE DIRECTION OF AN INDUCED EMF

Lenz's law states that an induced emf always gives rise to a current whose magnetic field will oppose the original change in flux.

Figure 4.3.2 applies the law to the relative motion between a magnet and a single coil of wire. Moving the magnet towards or away from the coil will induce an emf in the coil, as there is a change in flux. The induced emf will produce a current in the coil. And this induced current will then produce its own magnetic field. It is worth noting that Lenz's law is a necessary consequence of the law of conservation of energy: if Lenz's law were not true then the new magnetic field created by a changing flux would encourage that change, which would have the effect of adding energy to the universe.

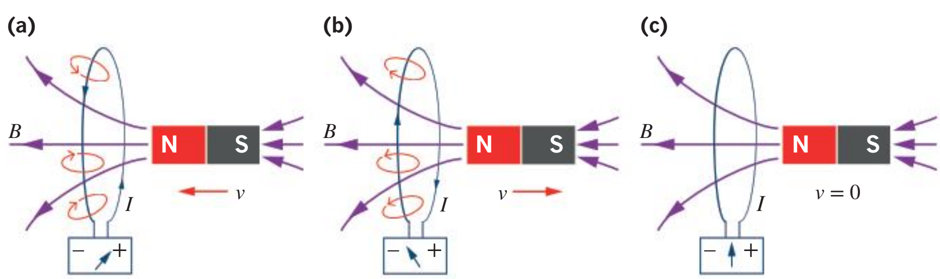


FIGURE 4.3.2 (a) The north end of a magnet is brought towards a coil from right to left, inducing a current that flows anticlockwise. (b) Pulling the north end of the magnet away from the coil from left to right induces a current in a clockwise direction. (c) Holding the magnet still creates no change in flux and hence no induced current.

Applying Lenz's law, the magnetic field created by the induced current will oppose the change in flux caused by the movement of the magnet. When the north end of a magnet is brought toward the loop from the right, the magnetic flux from right to left through the loop increases. The induced emf produces a current that flows anticlockwise around the loop when viewed from the right. The magnetic field created by this current, shown by the little circles around the wire, is directed from left to right through the loop. It opposes the magnetic field of the approaching magnet.

If the magnet is moved away from the loop, as in part (b) of Figure 4.3.2, the magnetic flux from right to left through the loop decreases. The induced emf produces a clockwise current when viewed from the right. This creates a magnetic field that is directed from right to left through the loop. This field is in the same direction as the original magnetic field of the retreating magnet. However, note that it is opposing the change in the magnet's flux through the loop by attempting to replace the declining flux.

And when the magnet is held stationary, as in part (c) of Figure 4.3.2, there is no change in flux to oppose and so no current is induced.

The right-hand grip rule and induced current direction

The right-hand grip rule (see Section 2.3 on p48) can be used to find the direction of the induced current. Keep in mind that the current must create a magnetic field that opposes the change in flux due to the relative motion of the magnet and conductor. Point your fingers through the loop in the direction of the field that is *opposing* the change and your thumb will then indicate the direction of the conventional current, as shown in Figure 4.3.3.

There are three distinct steps to determine the induced current direction according to Lenz's law:

1. What is the change that is happening?
2. What will *oppose* the change and/or restore the original conditions?
3. What must be the current direction to match this opposition?

These steps will be further examined using Worked example 4.3.1.

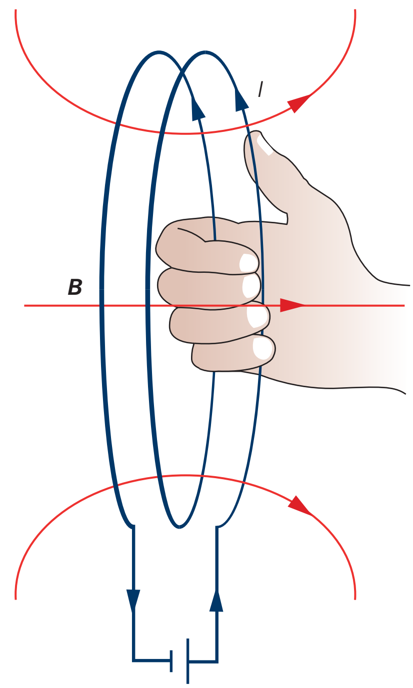
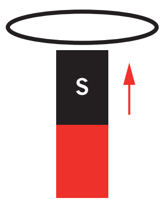


FIGURE 4.3.3 The right-hand grip rule can be used to determine the direction of a magnetic field from a current or vice versa. Your thumb points in the direction of the conventional current in the wire and your curled fingers indicate the direction of the magnetic field through the coil.

Worked example 4.3.1

INDUCED CURRENT IN A COIL FROM A PERMANENT MAGNET

The south pole of a magnet is brought upwards toward a horizontal coil initially held above it. In which direction will the induced current flow in the coil?

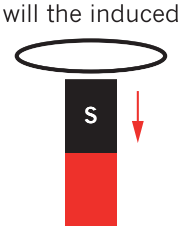


Thinking	Working
Consider the direction of the change in magnetic flux.	The magnetic field direction from the magnet will be downwards towards the south pole. The downwards flux from the magnet will increase as the magnet is brought closer to the coil. So the change in flux is increasing downwards.
What will oppose the change in flux?	The induced magnetic field that opposes the change would act upwards.
Determine the direction of the induced current required to oppose the change.	In order to oppose the change, the current direction would be anticlockwise when viewed from above (using the right-hand grip rule).

Worked example: Try yourself 4.3.1

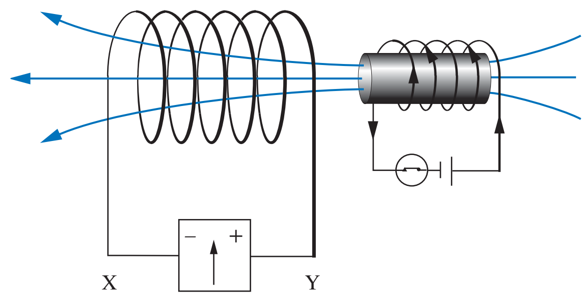
INDUCED CURRENT IN A COIL FROM A PERMANENT MAGNET

The south pole of a magnet is moved downwards away from a horizontal coil held above it. In which direction will the induced current flow in the coil?



Worked example 4.3.2

INDUCED CURRENT IN A COIL FROM AN ELECTROMAGNET

Instead of using a permanent magnet to change the flux in the loop in Worked example 4.3.1, an electromagnet (on the right, in the diagram below) could be used. What is the direction of the current induced in the solenoid when the electromagnet is: **(i)** switched on **(ii)** left on **(iii)** switched off?  	
Thinking	**Working**
Consider the direction of the change in magnetic flux for each case.	**(i)** Initially there is no magnetic flux through the solenoid. When the electromagnet is switched on, the electromagnet creates a magnetic field directed to the left. So the change in flux through the solenoid is increasing to the left. **(ii)** While the current in the electromagnet is steady, the magnetic flux through the solenoid is constant and the flux is not changing. **(iii)** In this case, initially there is a magnetic flux through the solenoid from the electromagnet directed to the left. When the electromagnet is switched off, there is no longer a magnetic flux through the solenoid. So the change in flux through the solenoid is decreasing to the left.
What will oppose the change in flux for each case?	**(i)** The magnetic field that opposes the change in flux through the solenoid is directed to the right. **(ii)** There is no change in flux and so there will be no opposition needed and no magnetic field created by the solenoid. **(iii)** The magnetic field that opposes the change in flux through the solenoid is directed to the left.
Determine the direction of the induced current required to oppose the change for each case.	**(i)** In order to oppose the change, the current will flow through the solenoid in the direction from X to Y (or through the meter from Y to X), using the right-hand grip rule. **(ii)** There will be no induced emf or current in the solenoid. **(iii)** In order to oppose the change, the current will flow through the solenoid in the direction from Y to X (or through the meter from X to Y), using the right-hand grip rule.

Worked example: Try yourself 4.3.2

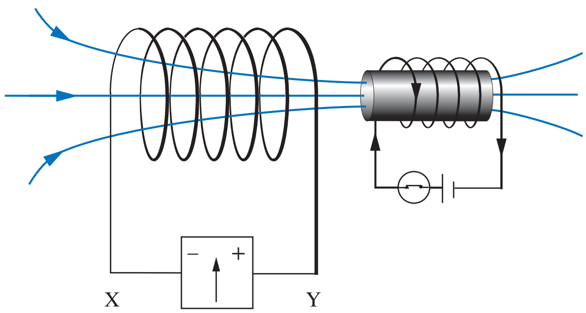
INDUCED CURRENT IN A COIL FROM AN ELECTROMAGNET

What is the direction of the current induced in the solenoid when the electromagnet is:

(i) switched on

(ii) left on

(iii) switched off?



INDUCED CURRENT DIRECTION BY CHANGING AREA

It's very important to note that an induced emf is created while there is a change in flux, no matter how that change is created. As magnetic flux $\Phi_B = B_\perp A$, a change can be created by any method that causes a relative change in the strength of the magnetic field, B, and/or the area of the coil perpendicular to the magnetic field. So an induced emf can be created in three ways:

- by changing the strength of the magnetic field
- by changing the area of the coil within the magnetic field
- by changing the orientation of the coil with respect to the direction of the magnetic field.

Figure 4.3.4 illustrates an example of the direction of an induced current that results during a decrease in the area of a coil.

As the area of the coil decreases due to its changing shape, the flux through the coil (which is directed into the page) also decreases. Applying Lenz's law, the direction of the induced current would oppose this change and will be such that it acts to increase the magnetic flux through the coil into the page. Using the right-hand grip rule, a current would therefore flow in a clockwise direction while the area is changing.

In Figure 4.3.5, the coil is being rotated within the magnetic field. The effect is the same as reducing the area. The amount of flux flowing through the coil is reduced as the coil changes from being perpendicular to the field to being parallel to the field. An induced emf would be created while the coil is being rotated. This becomes particularly important in determining the current direction in a generator.

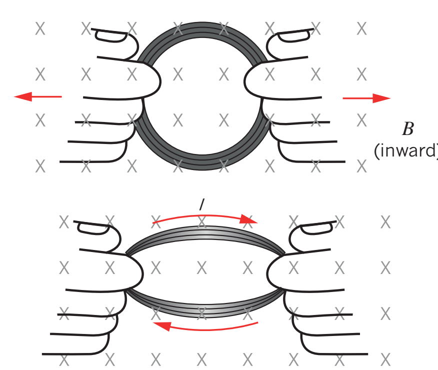


FIGURE 4.3.4 Inducing a current by changing the area of a coil. The amount of flux (the number of field lines) through the coil is reduced and an emf is therefore induced during the time that the change is taking place. The current flows in a direction that creates a field to oppose the reduction in flux into the page.

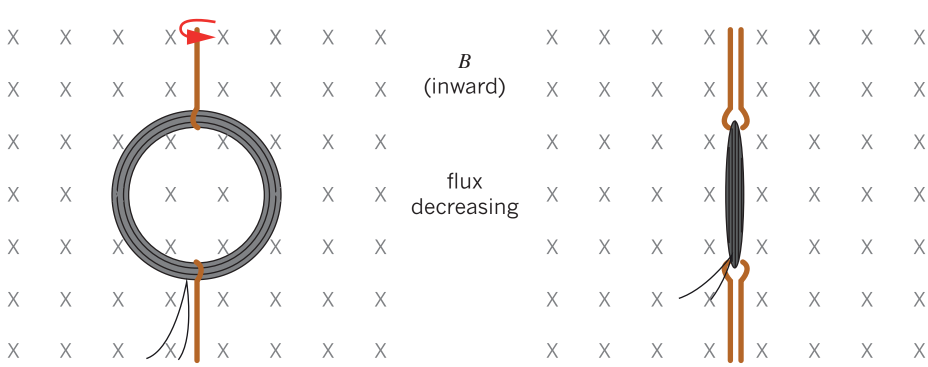
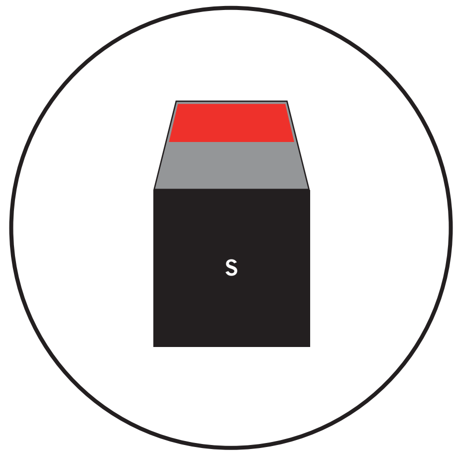


FIGURE 4.3.5 Changing the orientation of a coil within a magnetic field by rotating it reduces the amount of flux through the coil and so induces an emf in the coil while it is being rotated.

Worked example 4.3.3

FURTHER PRACTICE WITH LENZ'S LAW

The north pole of a magnet is moving towards a coil, into the page (the south pole is shown at the top looking down). In what direction will the induced current flow in the coil while the magnet is moving towards the coil?

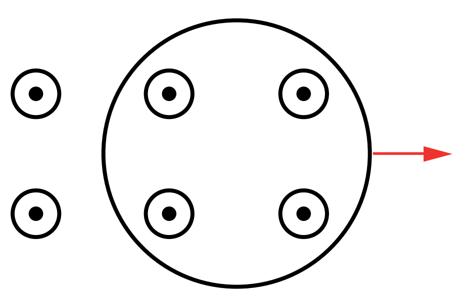


Thinking	Working
Consider the direction of the change in magnetic flux.	The magnetic field direction from the magnet will be away from the north pole, into the page. The flux from the magnet will increase as the magnet is brought closer to the coil. So the change in flux is increasing into the page.
What will oppose the change in flux?	The magnetic field that opposes the change would act out of the page.
Determine the direction of the induced current required to oppose the change.	In order to oppose the change, the current direction would be anticlockwise when viewed from above (using the right-hand grip rule).

Worked example: Try yourself 4.3.3

FURTHER PRACTICE WITH LENZ'S LAW

A coil is moved to the right and out of a magnetic field that is directed out of the page. In what direction will the induced current flow in the coil while the magnet is moving?



PHYSICSFILE

Eddy currents

Lenz's law is important for many practical applications such as metal detectors, induction stoves, and regenerative braking. These all rely on an eddy current, which is a circular electric current induced within a conductor by a changing magnetic field.

Applying Lenz's law, an eddy current will be in a direction that creates a magnetic field that opposes the change in magnetic flux that created it. Thus eddy currents can be used to apply a force that opposes the source of the motion of an external magnetic field. For example, if a metal plate is dragged out of a magnetic field, an eddy current will form within the plate that opposes the change in flux through the area of the plate, and thus opposes the motion of the plate itself due to the interaction of the magnetic fields (see Figure 4.3.6).

This is the basis of regenerative braking, where the drag of the opposing magnetic field is utilised as a braking force. An eddy current flowing through a conductor with some resistance will also lose energy to the conductor by heating it. This makes eddy currents useful for an induction stovetop, but a potentially major source of energy loss within an AC generator, motor, or transformer. Laminated cores with insulating material between the thin layers of iron are used in these applications to reduce the overall conductivity and suppress eddy currents.

The Earth's magnetic field is also a result of eddy currents. The energy that drives the Earth's dynamo comes from the enormous heat produced by radioactive decay deep in the Earth's core. The heat causes huge swirling convection currents of molten iron in the outer core. These convection currents of molten iron act rather like a spinning disk. They are moving in the Earth's magnetic field and so eddy currents are induced in them. It is these eddy currents that produce the Earth's magnetic field.

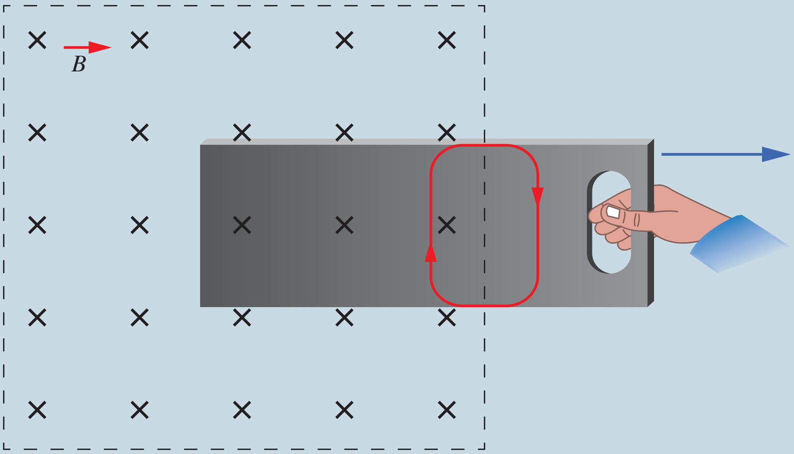


FIGURE 4.3.6 As the metal plate is moved towards the right, out of the magnetic field which is directed into the page, an eddy current forms in a clockwise direction. This eddy current would resist the motion of the plate.

PHYSICSFILE

Induction stoves

In contrast to a conventional gas or electric stove that heats via radiant heat from a hot source, an induction stove heats via the metal pot in which the food is being cooked. A coil of copper wire is placed within the cooktop (see Figure 4.3.7). The AC electricity supply produces a changing magnetic field in the coil. This induces an eddy current in the conductive metal pot. The resistance of the metal in the pot, in which the eddy current flows, transforms electrical energy into heat and cooks the food.

While induction cooktops have only reached the domestic market in relatively recent times, the first patents for induction cookers where issued in the early 1900s. They have significant advantages over traditional electric cooktops in that they allow instant control of cooking power (similar to gas burners), they lose less energy through ambient heat loss and heating time, and they have a lower risk of burn injuries. Overall, the heating efficiency of an induction cooktop is around 12% better than traditional electric cooktops and twice that of gas.

FIGURE 4.3.7 The coil of an induction zone within an induction cooktop. The large copper coil creates an alternating magnetic field.

PHYSICS IN ACTION

Superconductors and superconducting magnets

Technological breakthroughs have often led to advances in physics. This was the case in 1908 when Kamerlingh Onnes, at the University of Leiden in the Netherlands, succeeded in liquefying helium. Helium liquefies at 4.2 K (–268.9°C). It was known that the electrical resistance of metals decreases as they cool, so one of the first things that Onnes and his assistant did was to measure the resistance of some metals at these very low temperatures.

Onnes was hoping to find that as the temperature of mercury dropped towards absolute zero its resistance would also gradually drop towards zero. What they found, however, was a complete surprise. At 4.2 K its resistance vanished completely!

Onnes coined the word 'superconductivity' to describe this phenomenon. Soon he found that some other metals also became superconducting at extremely low temperatures: lead at 7.2 K and tin at 3.7 K, for example. Curiously, metals such as copper and gold, which are very good conductors at normal temperatures, do not become superconducting at all. Onnes was awarded the 1913 Nobel Prize in Physics for his work in low-temperature physics.

Much of the great promise of superconductivity has to do with the magnetic properties of superconductors. In a superconductor an induced current does not fade away! As the resultant field opposes the changing flux, the magnet is repelled. This gives rise to the 'magnetic levitation' effect that is by now well known (see Figure 4.3.8). On a large scale this could perhaps one day provide a virtually frictionless maglev (magnetic levitation) train.

Unfortunately, the superconducting metals lost their superconductivity in magnetic fields around 0.1 T, which is quite a moderate field. However, in the 1940s it was found that some alloys of elements such as niobium had higher 'critical temperatures' and, more particularly, retained their properties in stronger magnetic fields. By 1973 the niobium–germanium alloy Nb_3Ge held the record with a critical temperature of 23.2 K in a critical field of 38 T, an extremely strong field.

In 1986 an entirely new and exciting class of superconductors was discovered. Georg Bednorz and Karl Müller, working in Switzerland, found that compounds of some rare earth elements and copper oxide had considerably higher critical temperatures. They received the 1987 Nobel Prize in Physics for their work.

These new 'warm superconductors' are ceramic materials made by powdering and baking the metal compounds. Most ceramics are insulators; it was a combination of good science and inspired guesswork that led Müller to try such unlikely candidates for superconductivity. So far, the record is held by bismuth and thallium oxides with a critical temperature around 125 K—still rather cold, but significantly above the temperature of readily available liquid nitrogen (77 K).

Superconductivity, particularly in the newer materials, is still not fully understood. It can really only be discussed in terms of quantum physics, but one rather picturesque way of thinking about it is that electrons pair up and 'surf' electrical waves set up by each other in the crystal lattice of the material.

The promise of superconductivity is enormous: low friction transport, no-loss transmission of electricity, and smaller and more powerful electric motors and generators. Superconducting magnets might be used to contain the extremely hot plasma needed to bring about hydrogen fusion, producing almost pollution-free energy in much the same way that the Sun does. There are, however, many difficulties to be overcome before these promises can be realised.

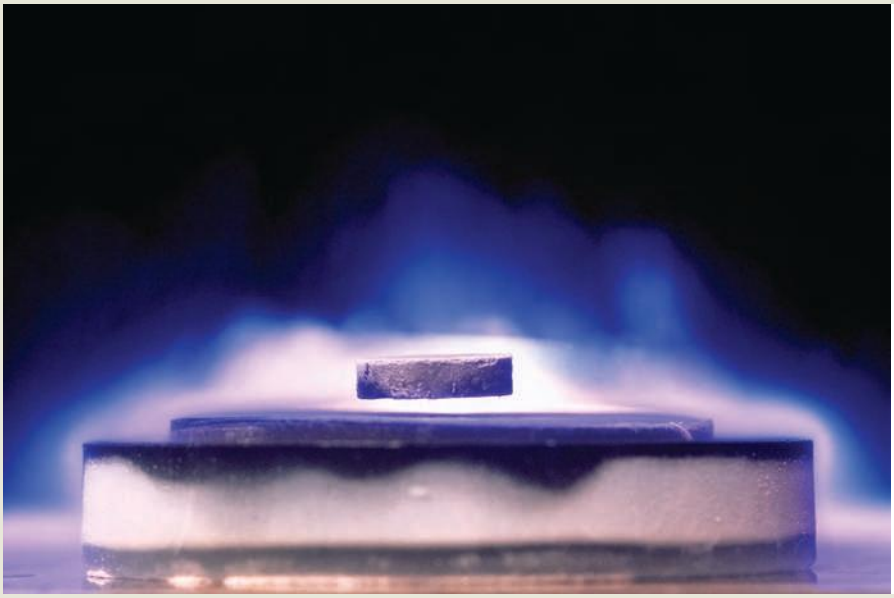


FIGURE 4.3.8 A disk magnet is repelled by a superconductor because the magnet induces a permanent current into the superconductor, which results in an opposing field.

ELECTRIC POWER GENERATORS

These days we take the supply of electric power to our homes, schools and businesses for granted, and yet it was only in 1881 that the Victorian Electric Light Company demonstrated an electric lamp lit by a gas-powered generator outside its premises in Swanston Street, Melbourne.

The electric **generator** is probably the most important practical application of Faraday's discovery of electromagnetic induction. The principle of electric power generators is the same whether the result is alternating current or direct current. Relative motion between a coil and a magnetic field induces an emf in the coil. In small generators, the coil is rotated within a magnetic field, but in large power stations, car alternators, and other industrial-level production, the coils are stationary and an electromagnet rotates inside them.

This might all sound quite similar to the way electric motors work (see Chapter 3). In fact, it is—a generator is basically just the inverse of a motor.

Induced emf in an alternator or generator

A basic electric generator, or **alternator**, consists of many coils of wire wound on an iron core framework. This is called an **armature** and it is made to rotate in a magnetic field. The axle is turned by some mechanical means—mechanical energy is being converted to electrical energy—and an emf is induced in the rotating coil.

Consider a single loop of wire in the generator shown in Figure 4.3.9. The loop is rotated clockwise in a uniform magnetic field, B. The amount of flux threading through the loop will vary as it rotates. It is this change in flux that induces the emf.

Lenz's law tells you that as the flux in the loop decreases from position (a) to (b) in Figure 4.3.9, the induced current will be in a direction such as to restore a magnetic field in the same direction, relative to the loop, as the external field. The right-hand grip rule can then be used to show that the direction of the induced current will be D → C → B → A.

The direction of the induced current will reverse every time the plane of the loop reaches a point perpendicular to the field. The magnitude of the induced emf will be determined by the rate at which the loop is rotating. It will be a maximum when the rate of change of flux is a maximum. This is when the loop has moved to a position parallel to the magnetic field and the flux through the loop is zero, i.e., the gradient of the flux versus time graph shown in Figure 4.3.10 is a maximum.

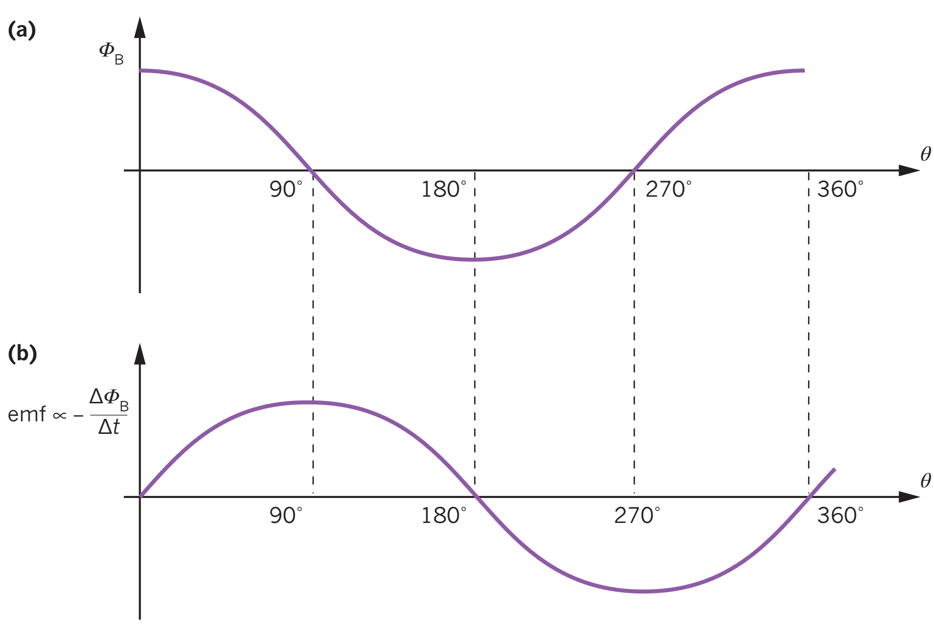


FIGURE 4.3.10 (a) The flux, Φ_B, through the loop of Figure 4.3.9 as a function of the angle between the field and the normal to the plane of the area, θ. (b) The rate of change of flux and hence emf through the loop as a function of the angle between the field and the normal to the plane of the loop, θ. The loop is rotating at a constant speed.

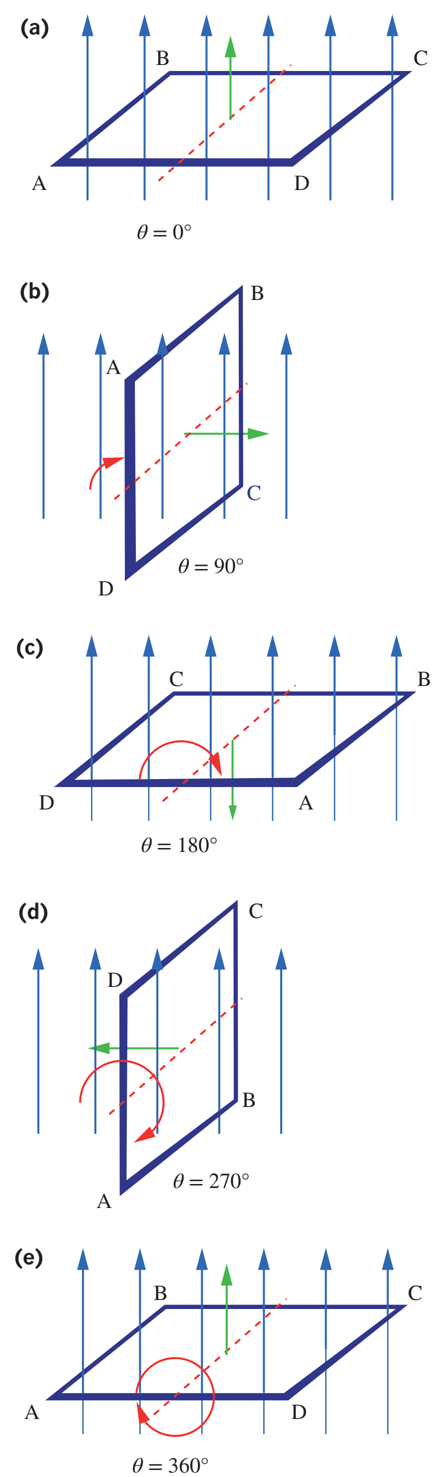


FIGURE 4.3.9 A single loop of a generator rotating in a magnetic field. (a) The plane of the area of the loop is perpendicular to the field B and the amount of flux $\Phi = BA$ is at a maximum. (b) The loop has turned one quarter of a turn and is parallel to the field; $\Phi = 0$. (c) As the loop continues to turn, the flux increases to a maximum but in the opposite sense relative to the loop in (a); $\Phi = -BA$. (d) The flux then decreases to zero again as the loop is parallel to the field before repeating the cycle again from (e) onwards.

An alternative way to think about how the emf changes as the loop rotates is to remember that the emf is actually created as the wires AB and CD cut across the magnetic field lines. Maximum emf occurs when these wires cut the magnetic field lines perpendicularly, when θ is 90° or 270°, and zero emf occurs when the motion of these wires is parallel to the field lines when θ is 0°, 180° or 360°.

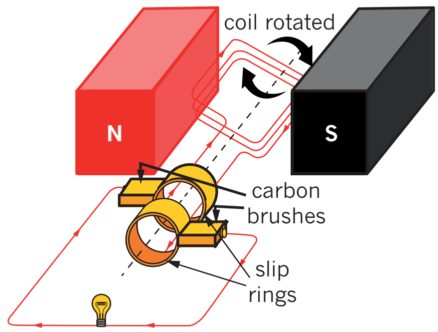


FIGURE 4.3.11 A schematic of an AC generator showing the main features.

AC generators and alternators

A generator's construction is basically the same as a motor. The main components of an AC generator are shown in Figure 4.3.11.

Consider a coil, or armature, with a number of turns, being rotated in a magnetic field, inducing an emf as shown previously in Figure 4.3.9. The resultant emf alternates in direction as shown by the graph going above and below the zero emf line in Figure 4.3.10. This type of emf or voltage produces an alternating current (AC) in the coil. How this alternating current in the coil is harnessed determines if the device is an AC alternator or a DC generator.

As was stated earlier, many industrial generators will instead have the coils remain static and the electromagnet rotates within them. The principle of inducing an emf is the same. The coil itself may take a variety of shapes, sizes and positions.

If the output from the coils is transferred to a circuit via continuous **slip rings**, the alternating current in the coil will be maintained at the output. The slip rings also allow the coil to rotate without tangling. Carbon **brushes** press against the slip rings to allow a constant output to be transferred to a circuit without a fixed point of connection.

PHYSICS IN ACTION

Three-phase generators

Many industrial applications require a more constant maximum voltage than is possible from a single coil. These applications require a three-phase power supply. The coils are arranged such that the emfs vary at the same frequency, but with the peaks and troughs of their waveforms offset to provide three complementary currents with a phase separation of one-third of a cycle, or 120°. The resulting output of all three phases maintains an emf near the maximum voltage more continuously. Standard electrical supplies include three phases, but most home applications only require a single phase to be connected.

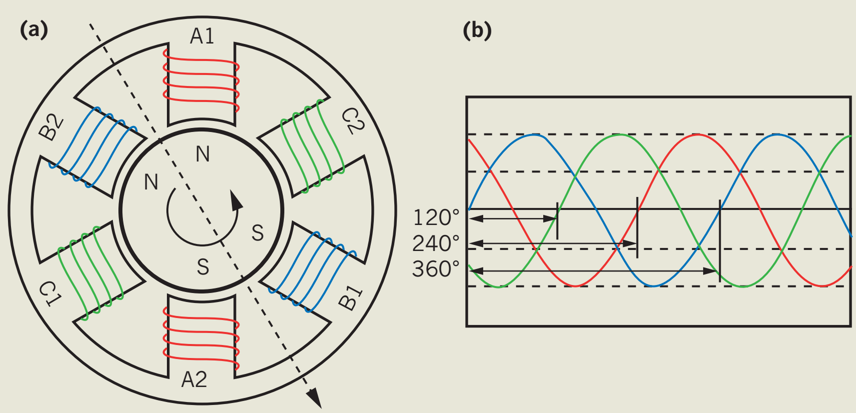


FIGURE 4.3.12 (a) A three-phase power supply has three coils, each producing an output 120° out of phase with the adjoining coil. (b) The resulting output can be combined for a more constant supply voltage.

DC generators

A DC generator is much like an AC generator in basic design. The continuous slip rings are replaced by a **split ring commutator**. That is, the ring picking up the output from the coils has two breaks (or splits) in it at opposite sides of the ring. The direction of the output is changed by the commutator every half turn so that the output current is always in the same direction (see Figure 4.3.13(a)). The output will still vary from zero to a maximum every half cycle. The output can be smoothed by placing a capacitor in parallel with the output. More commonly, the use of multiple armature windings and more splits in the split ring commutator can smooth the output by ensuring that the output is always connected to an armature that is in the position for generating maximum emf (see Figure 4.3.13(b)).

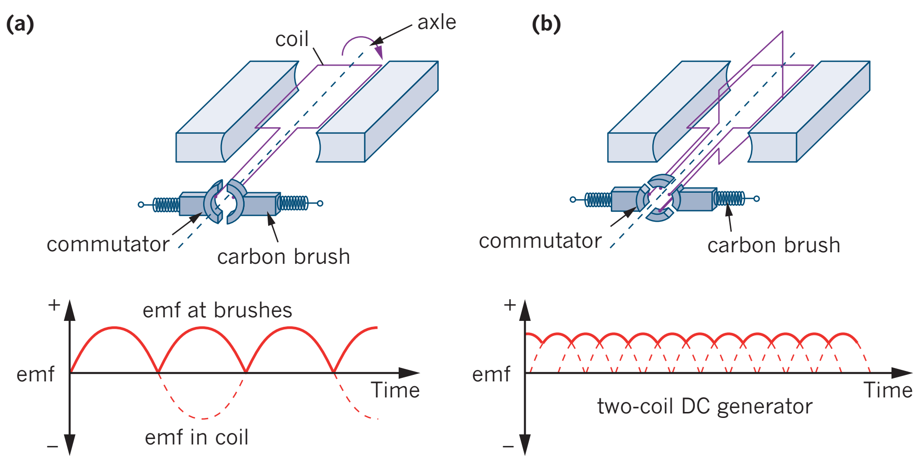


FIGURE 4.3.13 (a) A DC generator has a commutator to reverse the direction of the alternating current every half cycle and so produce a DC output. (b) Multiple armature windings can smooth the output.

In the past, cars used DC generators to power ancillary equipment. More common now is the use of AC generators or alternators, which avoid the problems of wear and sparking across the commutator inherent in the design of DC generators by using a moving electromagnet inside a set of stationary coils to generate current.

PHYSICS IN ACTION

Back emf in DC motors

The description of the construction and operation of a generator shows that a DC motor and a generator share a lot in common and may even function either way. In fact, every motor can also be used as a generator. The motors of electric trains, for instance, work as generators when a train is slowing down, converting kinetic energy to electrical energy and putting it back into the electrical supply grid. Regenerative brakes in cars work in a similar way. A DC motor will also generate an emf when running normally. This is termed the 'back emf'.

The back emf generated in a DC motor is the result of current produced in response to the rotation of the rotor inside the motor in the presence of an external magnetic field. The back emf, following Lenz's law, opposes the change in magnetic flux that created it, so this induced emf will be in the opposite direction to the emf creating it. The net emf used by the motor is thus always less than the supplied voltage:

$\varepsilon_{net} = V - \varepsilon_{back}$

As the motor increases speed, the current induced in it will increase and the back emf will also increase. When a load is applied to the motor, the speed will generally reduce. This will reduce the back emf and increase the current in the motor. If the load brings the motor to a sudden halt—say, an electric drill bit getting stuck—the current may be high enough to burn out the motor and the motor windings. To protect the motor, a resistor is placed in series. It is switched out of the circuit when the current drops below a predetermined level and is switched back into the circuit for protection once the level is exceeded.

ALTERNATING VOLTAGE AND CURRENT

An AC generator produces an alternating current varying sinusoidally over time with the change in magnetic flux. The maximum emf is only achieved for particular points in time. In Australia, mains power oscillates at 50 Hz and reaches a peak voltage of ±340 V each cycle or a peak-to-peak voltage of 680 V (see Figure 4.3.14).

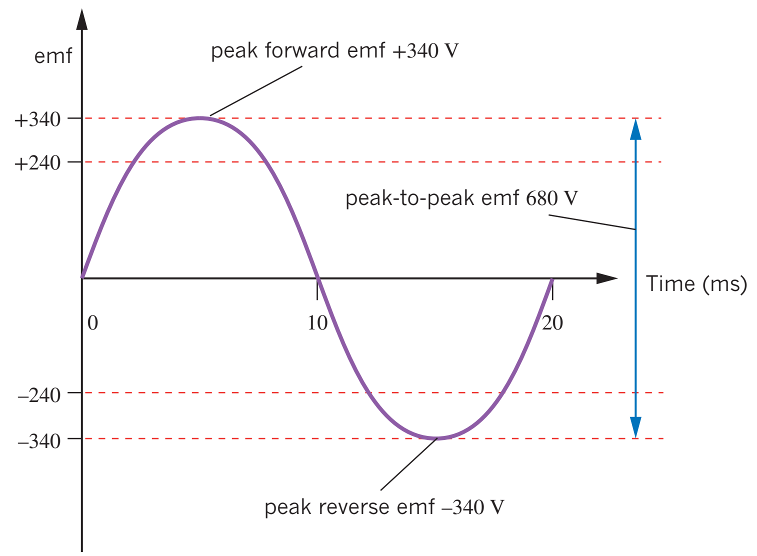


FIGURE 4.3.14 The voltage in Australian power points oscillates between +340 V and −340 V, 50 times each second. The value of a DC supply that would supply the same average power is 240 V.

It is often more useful to know the average power produced in a circuit. The average power can be obtained by using a value for the voltage and current equal to the peak values divided by $\sqrt{2}$. This is referred to as the **root mean square** or rms value.

EXTENSION

Deriving the root mean square formulae

In an AC circuit, the power produced in a resistor is equal to: $\frac{V^2}{R}\sin^2\theta$

The average power will be given by:

$$\frac{1}{2}\frac{V_p^{\,2}}{R}$$

If this same power was to be supplied by a steady (DC) source, the voltage V_{ave} of this source would have to be such that:

$$\frac{V_{ave}^{\,2}}{R} = \frac{1}{2}\frac{V_p^{\,2}}{R}$$

Simplifying:

$$V_{ave}^{\,2} = \frac{V_p^{\,2}}{2}$$

$$V_{ave} = \frac{V_p}{\sqrt{2}}$$

This voltage is known as the root mean square voltage or V_{rms}. It is the value of a steady voltage that would produce the same power as an alternating voltage with a peak value equal to $\sqrt{2}$ times as much.

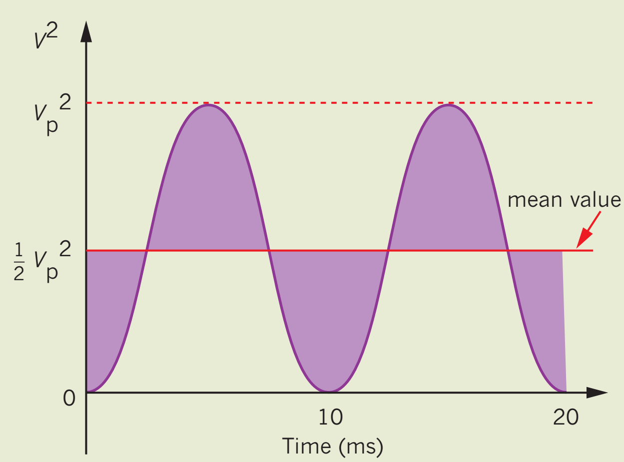


FIGURE 4.3.15 The power transmitted is proportional to the area under a V^2 graph. Clearly, the power transmitted by an AC circuit (with V_p) is the same as that in a DC circuit with a voltage equal to the square root of $\frac{1}{2}(V_p)^2$, that is $\frac{V_p}{\sqrt{2}}$.

In effect, the rms values are the values of a DC supply that would be needed to provide the same average power as the AC supply. It is the rms value of the voltage $\left(\frac{340}{\sqrt{2}} = 240\right)$ that is normally quoted. This is the effective average value of the voltage and is the value that should be used to find the actual power supplied each cycle by an AC supply. So:

$$V_{rms} = \frac{V_p}{\sqrt{2}}$$

$$I_{rms} = \frac{I_p}{\sqrt{2}}$$

$$P_{rms} = V_{rms} \times I_{rms} = \frac{1}{2} V_p I_p, \text{ and}$$

$$P_p = \sqrt{2}V_{rms} \times \sqrt{2}I_{rms} = 2V_{rms}I_{rms}$$

Worked example 4.3.4

PEAK AND RMS AC CURRENT VALUES

A 60 W light globe is connected to a 240 V AC circuit. What is the peak power use of the light globe?	
Thinking	**Working**
Note that the values given in the question represent rms values. Power is $P = VI$ so both V and I must be known to calculate the power use. The voltage V is given, and the current I can be calculated from the rms power supplied.	$P_{rms} = V_{rms}I_{rms}$ $I_{rms} = \frac{P_{rms}}{V_{rms}}$ $= \frac{60}{240} = 0.25$ A
Substitute in known quantities and solve for peak power.	$P_p = \sqrt{2}V_{rms} \times \sqrt{2}I_{rms} = 2V_{rms}I_{rms}$ $= 2 \times V_{rms} \times I_{rms}$ $= 2 \times 240 \times 0.25$ $= 120$ W

Worked example: Try yourself 4.3.4

PEAK AND RMS AC CURRENT VALUES

A 1000 W kettle is connected to a 240 V AC power outlet. What is the peak power use of the kettle?

4.3 Review

SUMMARY

- An induced emf always gives rise to a current whose magnetic field will oppose the original change in flux.
- There are three distinct steps to determine the induced current direction according to Lenz's law:
  1. What is the change that is happening?
  2. What will oppose the change and/or restore the original conditions?
  3. What must be the current direction to match this opposition?
- An induced emf can be created in three ways:
  - by changing the strength of the magnetic field
  - by changing the area of the coil within the magnetic field
  - by changing the orientation of the coil with respect to the direction of the magnetic field.
- The principle of electric power generators is the same whether the result is alternating current or direct current. Relative motion between a coil and a magnetic field induces an emf in the coil.
- A generator's, or alternator's, construction is very similar to that of an electric motor.
- A coil rotated in a magnetic field will produce an alternating induced current in the coil. How that current is harnessed will determine if the device is an AC alternator or a DC generator.
- An AC alternator has slip rings that transfer the alternating nature of the current in the coil to the output. A DC generator has a split ring commutator to reverse the current direction every half turn so that the output current is always in the same direction.
- The alternating current produced by power stations and supplied to cities varies sinusoidally at a frequency of 50 Hz. The peak value of the voltage of domestic power (V_p) is ±340 V, and the peak-to-peak voltage (V_{p-p}) is 680 V.
- The root mean square voltage, V_{rms}, is the value of an equivalent steady voltage (DC) supply that would provide the same power.

  $V_{rms} = \frac{V_p}{\sqrt{2}}$
- The rms value of domestic mains voltage in Australia is 240 V.
- The average power in a resistive AC circuit is:

  $P = V_{rms}I_{rms}$

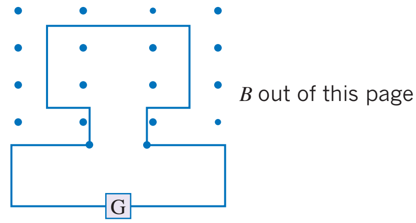
  $= \frac{1}{2} \times V_p \times I_p$

KEY QUESTIONS

1 A conducting loop is located in an external magnetic field whose direction (but not necessarily magnitude) remains constant. A current is induced in the loop. Which of the following alternatives best describes the direction of the magnetic field due to the induced current?

A It will always be in the same direction as the external magnetic field.
B It will always be in the opposite direction to the external magnetic field.
C It will be in the same direction as the external magnetic field if the external magnetic field gets weaker, and it will be in the opposite direction to the external magnetic field if the external magnetic field gets stronger.
D The direction can't be determined from the information supplied.

2 A rectangular conducting loop forms the circuit shown below. The plane of the loop is perpendicular to an external magnetic field whose magnitude and direction can be varied. The initial direction of the field is out of the page.



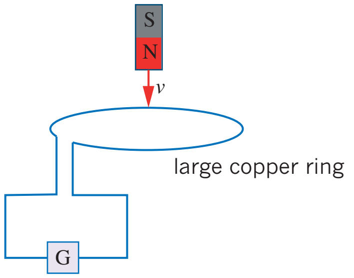
a When the magnetic field is switched off, what will be the direction of the magnetic field due to the induced current?

A out of the page
B into the page
C clockwise
D anticlockwise
E left to right
F right to left

b When the direction of the magnetic field is reversed, what is the direction of the magnetic field due to the induced current?

A out of the page
B into the page
C clockwise
D anticlockwise
E left to right
F right to left

3 A bar magnet is falling towards the centre of a horizontal copper coil, as shown below.



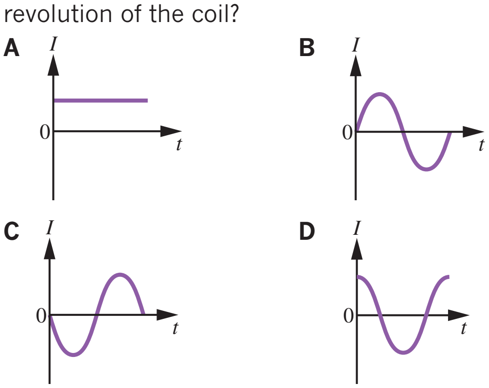
a What is the direction (as seen from above) of the induced current in the coil when the magnet is in the position shown in the diagram?

b Name four factors that would influence the magnitude of the induced current in the copper ring.

4 The back emf generated in a DC motor is the result of current produced in response to the rotation of the armature in the motor in the presence of an external magnetic field. As a result of the back emf, what will the net emf used by a DC motor be?

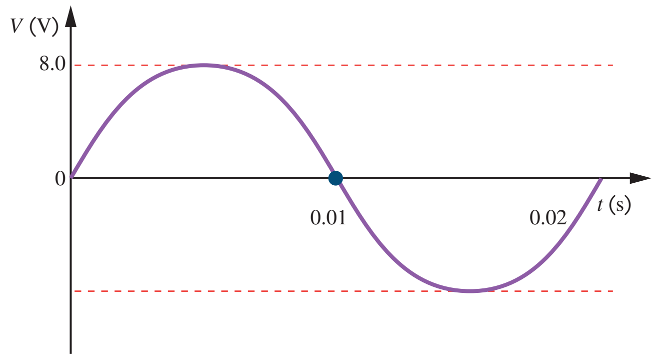
A the same as the supplied voltage
B less than the supplied voltage
C greater than the supplied voltage
D greater or less than the supplied voltage, depending on the speed of the motor

5 Assuming that an anticlockwise rotation of a coil starting from $\theta = 0°$ perpendicular to a constant magnetic field initially produces a positive current, which of the graphs best illustrates the variation of the induced current as a function of time for one full revolution of the coil?



The following information relates to questions 6 and 7.

A simple generator consists of a coil of $N = 1000$ turns, each of radius 10 cm, mounted on an axis in a uniform magnetic field of strength, B. The following graph shows the voltage output as a function of time when the coil is rotated at a frequency of 50 Hz.

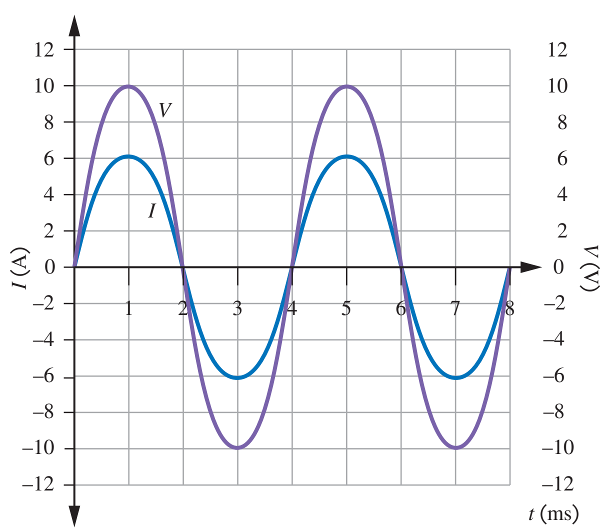


6 Determine the values of V_p, $V_{p\text{–}p}$ and V_{rms}.

7 The generator is modified so that the magnetic field strength is doubled and the frequency of rotation is increased to 100 Hz. The radius of the coil is halved to 5.0 cm. Draw a line graph representing the new output from the generator.

8 An AC supply of frequency 50 Hz is connected to a circuit, resulting in an rms current of 1.0 A being observed. Draw a graph that shows one full period of the variation of current with time for this circuit.

9 A student decides to test the output power of a new amplifier by using a voltage sensor to capture and display the alternating current I and voltage V that it produces. The result is shown below.



What is the rms power rating of the amplifier?

10 An electric toaster designed to operate at a V_{rms} of 240 V has a power rating of 600 W. What is the peak current in the heating element?

4.4 Supplying electricity—transformers and large-scale power distribution

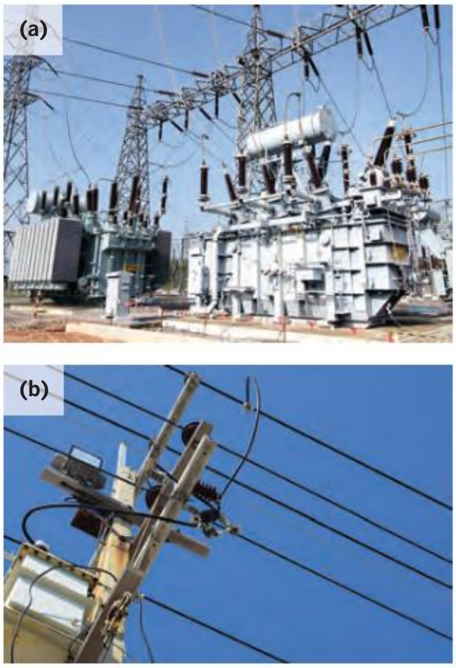


FIGURE 4.4.1 (a) View of transformers at an electrical substation. The substation takes electricity from the distribution grid and converts it to lower voltages used by industrial or residential equipment. More common are the smaller distribution transformers, found on every suburban street (b). See if you can locate at least one on your street.

When Faraday first discovered electromagnetic induction, he had effectively invented the transformer. A transformer is a device for increasing and decreasing an AC voltage. Transformers can be found in just about any electrical device, are an essential part of any electrical distribution system and are the focus of this section (see Figure 4.4.1).

THE WORKINGS OF A TRANSFORMER

A **transformer** works on the principle of a changing magnetic flux inducing an emf. No matter what the size or application, a transformer will consist of two coils known as the *primary* and *secondary* coils. The changing flux originates with the alternating current supplied to the primary coil. The changing magnetic flux is directed to the secondary coil where the changing flux will induce an emf in that coil (see Figure 4.4.2).

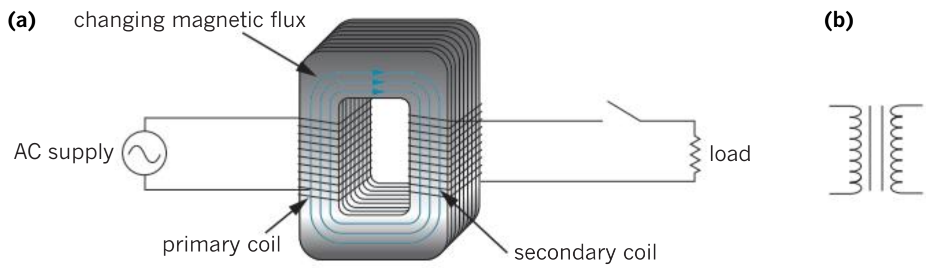


FIGURE 4.4.2 (a) In an ideal transformer, the iron core ensures that all the flux generated in the primary coil also passes through the secondary coil. (b) The symbol used in circuit diagrams for an iron-core transformer.

The two coils can be interwoven using insulated wire or they can be linked by a soft iron core, laminated to minimise eddy current losses. Transformers are designed so that nearly all of the magnetic flux produced by the primary coil will pass through the secondary coil. In an **ideal transformer** the assumption is that this will be 100% efficient and energy losses can be ignored. In a real transformer, this assumption remains a good approximation. Transformers are one of the most efficient devices around, with practical efficiencies often being better than 99%.

PHYSICSFILE

Laminations

Eddy currents that are set up in the iron core of transformers can generate a considerable amount of heat. Energy that has been lost from the electrical circuit and the transformer as heat may become a fire hazard. To reduce eddy current losses, the core is made of laminations, which are thin plates of iron electrically insulated from each other and placed so that the insulation between the laminations interrupts the eddy currents.

AC VERSUS DC

The power distribution system works on alternating current. That may seem odd when many devices run on direct current, but one of the primary reasons is the ease with which alternating current can be transformed from one voltage to another.

A transformer works on the basis of a changing current in the primary coil inducing a changing magnetic flux. This in turn induces a current in the secondary coil. For this to work, the original current must be constantly changing, as it does in an AC supply.

A DC voltage has a constant, unchanging current. With no change in the size of the current, no changing magnetic flux will be created by the primary coil and, hence, no current is induced in the secondary coil. Transformers do not work with the constant current of a DC electrical supply. There will be a very brief induced current when a DC supply is turned on, and a change occurs from zero current to the supply level. There is a similar spike if the DC supply is switched off, but while the DC supply is constant there is no change in magnetic flux to induce a current in the secondary coil.

THE TRANSFORMER EQUATION

When an AC voltage is connected to the primary coil of a transformer, the changing magnetic field will induce an AC voltage of the same frequency as the original supply in the secondary coil. The voltage in the secondary coil will be different and depends upon the number of turns in each coil.

From Faraday's law, the average voltage in the primary coil, V_p, will affect the rate at which the magnetic flux changes:

$$V_1 = N_1 \frac{\Delta \Phi_B}{\Delta t}$$

or

$$\frac{\Delta \Phi_B}{\Delta t} = \frac{V_1}{N_1}$$

where N_p is the number of turns in the primary coil.

The induced voltage in the secondary coil, V_s, will be

$$V_2 = N_2 \frac{\Delta \Phi_B}{\Delta t}$$

and

$$\frac{\Delta \Phi_B}{\Delta t} = \frac{V_2}{N_2}$$

where N_s is the number of turns in the secondary coil.

Assuming that there is little or no loss of flux between the primary and secondary coil, then the flux in each will be the same and

$$\frac{V_1}{N_1} = \frac{V_2}{N_2}$$

or

$$\frac{V_2}{V_1} = \frac{N_2}{N_1}$$

The transformer equation, relating voltage and number of turns in each coil, is:

$$\frac{V_1}{V_2} = \frac{N_1}{N_2} \text{ or } \frac{V_2}{V_1} = \frac{N_2}{N_1} \text{ or } \frac{V_1}{N_1} = \frac{V_2}{N_2}$$

where the subscript '1' refers to the primary or first coil, and the subscript '2' refers to the secondary coil.

The transformer equation explains how the secondary (output) voltage is related to the primary input voltage. Either the rms voltage for both or the peak voltage for both can be used.

A **step-up transformer** increases the secondary voltage compared with the primary voltage. The secondary voltage is greater than the primary voltage and the number of turns in the secondary coil is greater than the number of turns in the primary coil, i.e. if $N_2 > N_1$ then $V_2 > V_1$.

A **step-down transformer** decreases the secondary voltage compared with the primary voltage. The secondary voltage is less than the primary voltage and the number of turns in the secondary coil is less than the number of turns in the primary coil, i.e. if $N_2 < N_1$ then $V_2 < V_1$.

Worked example 4.4.1

TRANSFORMER EQUATION—VOLTAGE

A transformer is built into a portable radio to reduce the 240 V supply voltage to the required 12 V for the radio. If the number of turns in the secondary coil is 100, what is the number of turns required in the primary coil?

Thinking	Working
State the relevant quantities given in the question. Choose a form of the transformer equation with the unknown quantity in the top left position.	$V_2 = 12$ V $V_1 = 240$ V $N_2 = 100$ turns $N_1 = ?$ $\frac{N_1}{N_2} = \frac{V_1}{V_2}$
Substitute the quantities into the equation, rearrange and solve for N_1.	$\frac{N_1}{100} = \frac{240}{12}$ $N_1 = \frac{100 \times 240}{12}$ $= 2000$ turns

Worked example: Try yourself 4.4.1

TRANSFORMER EQUATION—VOLTAGE

A transformer is built into a phone charger to reduce the 240 V supply voltage to the required 6 V for the charger. If the number of turns in the secondary coil is 100, what is the number of turns required in the primary coil?

POWER OUTPUT

Although a transformer very effectively increases or decreases an AC voltage, energy conservation means that the output power cannot be any greater than the input power. Since a well-designed transformer with a laminated core can be more than 99% efficient, the power input can be considered equal to the power output, making it an 'ideal' transformer.

Since power supplied is $P = VI$, then:

$$V_1 I_1 = V_2 I_2$$

The transformer equation can then be written in terms of current, I.

The transformer equation, relating current and the number of turns in each coil:

$$\frac{I_1}{I_2} = \frac{N_2}{N_1} \text{ or } \frac{I_2}{I_1} = \frac{N_1}{N_2} \text{ or } \frac{I_1}{N_2} = \frac{I_2}{N_1}$$

Note carefully that the number-of-turns ratio for currents is the *inverse* of that for the transformer equation written in terms of voltage.

PHYSICSFILE

Overload

A transformer will be overloaded if too much current is drawn and the resistive power loss in the wires becomes too great. There will be a point at which the transformer starts to overheat rapidly. For this reason, it is important not to exceed the rated capacity of a transformer.

PHYSICSFILE

Standby power

Because very little current will flow in the primary coil of a good transformer to which there is no load connected, it will use little power when not in use. However, this 'standby power' can add up to around 10% of power use. This is why devices such as TVs and computers should be switched completely off when not in use. Over the whole community, standby power amounts to megawatts of wasted power and unnecessary greenhouse emissions! Special switches, such as the 'Ecoswitch' shown below, have been developed that can be connected between the power outlet and the device to make it easier to remember to turn devices completely off when not in use.

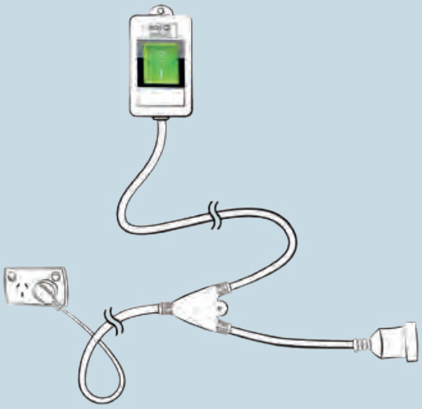


FIGURE 4.4.3 Standby switches such as the 'Ecoswitch' make it easier and more convenient to turn devices completely off when not in use, saving up to 10% on power bills.

Worked example 4.4.2

TRANSFORMER EQUATION—CURRENT

A radio with 2000 turns in the primary coil and 100 turns in its secondary coil draws a current of 4.0 A. What is the current in the primary coil?	
Thinking	**Working**
State the relevant quantities given in the question. Choose a form of the transformer equation with the unknown quantity in the top left position.	$I_2 = 4.0$ A $N_2 = 100$ turns $N_1 = 2000$ turns $I_1 = ?$ $\frac{I_1}{I_2} = \frac{N_2}{N_1}$
Substitute the quantities into the equation, rearrange and solve for I_1.	$\frac{I_1}{4.0} = \frac{100}{2000}$ $I_1 = \frac{4.0 \times 100}{2000}$ $= 0.20$ A

Worked example: Try yourself 4.4.2

TRANSFORMER EQUATION—CURRENT

A phone charger with 4000 turns in the primary coil and 100 turns in its secondary coil draws a current of 0.50 A. What is the current in the primary coil?

Worked example 4.4.3

TRANSFORMERS—POWER

The power drawn from the secondary coil of the transformer by a portable radio is 48 W. What power is drawn from the mains supply if the transformer is an ideal transformer?	
Thinking	**Working**
The energy efficiency of a transformer can be assumed to be 100%. The power in the secondary coil will be the same as that in the primary coil.	The power drawn from the mains supply is the power in the primary coil, which will be the same as the power in the secondary coil: $P = 48$ W.

Worked example: Try yourself 4.4.3

TRANSFORMERS—POWER

The power drawn from the secondary coil of the transformer by a phone charger is 3 W. What power is drawn from the mains supply if the transformer is an ideal transformer?

POWER FOR CITIES: LARGE-SCALE AC SUPPLY

In your school experiments using electrical circuits it is likely that you have ignored the resistance of the connecting wires because the wires (generally made from copper) are good conductors, and so the resistance is very small over short distances. However, over large distances, even relatively good electrical conductors like copper have a significant resistance.

Modern cities use huge amounts of electrical energy, most of which is supplied from power stations built at a considerable distance from the metropolitan areas. The efficient transmission of the electrical energy with the least amount of power loss over that distance is therefore a very important consideration for electrical engineers, particularly given the vast distances between population centres in Australia.

The power lost in an electrical circuit is given by $P = \Delta VI$, where ΔV is the voltage drop across the load. Recalling Ohm's law, $\Delta V = IR$, and substituting, the power can be expressed in terms of either current and load resistance or voltage drop and load resistance:

$$P_{loss} = \Delta VI = I^2R = \frac{\Delta V^2}{R}$$

By considering the form of the equation including the current carried by the circuit and its electrical resistance ($P_{loss} = I^2R$), it is clear that transmitting large amounts of power using a large current will create very large power losses. If the current in the power lines can be reduced, it will significantly reduce the power loss. Since the power loss is proportional to the square of the current then if the current is reduced by a factor of 3, for example, the power loss will be reduced by a factor of 3^2 or 9.

The challenge, then, is to transmit the large amounts of power being produced at power stations using a very low current. Transformers are the most common solution to this problem. Using a step-up transformer near the power station, the voltage is increased by a certain factor and, importantly, the current is decreased by the same factor. Due to the $P_{loss} = I^2R$ equation, the power lost during transmission is reduced by the square of that factor.

At this point you might be confused by the alternative equation for power loss: $P_{loss} = \frac{\Delta V^2}{R}$. A simple misunderstanding could make you think that increasing the voltage through the use of a step-up transformer would actually lead to greater power loss, if you use this equation to calculate power loss. However, ΔV represents the voltage drop in a circuit. You must be careful not to confuse the voltage being *transmitted along* the wires with the voltage *drop across* the wires. So, even though the voltage being transmitted is increased through the use of a step-up transformer, the voltage drop across the wires would be reduced since $\Delta V = IR$, and thus the power loss would also be reduced.

AC power from the generator is readily stepped up by a transformer to somewhere between 240 kV and 500 kV prior to transmission. Once the electrical lines reach the city, the voltage is stepped down in stages at electrical substations for distribution. The power lines in streets will have a voltage of around 2400 V, before being stepped down via small distribution transformers to 240 V for home use.

Worked example 4.4.4

TRANSMISSION-LINE POWER LOSS

300 MW is to be transmitted from the Hazelwood power station to Melbourne along a transmission line with a total resistance of 1.0 Ω. What would be the total transmission power loss if the initial voltage along the line was 250 kV?	
Thinking	**Working**
Convert the values to SI units.	$P = 300\text{ MW} = 300 \times 10^6\text{ W}$ $V = 250\text{ kV} = 250 \times 10^3\text{ V}$
Determine the current in the line based on the required voltage.	$P = VI \therefore I = \frac{P}{V}$ $I = \frac{300 \times 10^6}{250 \times 10^3}$ $= 1200\text{ A}$
Determine the corresponding power loss.	$P = I^2R$ $= 1200^2 \times 1$ $= 1.44 \times 10^6\text{ W}$ or 1.44 MW

Worked example: Try yourself 4.4.4

TRANSMISSION-LINE POWER LOSS

300 MW is to be transmitted from the Hazelwood power station to Melbourne along a transmission line with a total resistance of 1.0 Ω. What would be the total transmission power loss if the voltage along the line was now to be 500 kV?

Worked example 4.4.5

VOLTAGE DROP ALONG A TRANSMISSION LINE

Power is to be transmitted from the Loy Yang power station to Melbourne along a transmission line with a total resistance of 1.0 Ω. The current is 1200 A. What voltage would be needed at the Loy Yang end of the transmission line to achieve a supply voltage of 250 kV?

Thinking	Working
Determine the voltage drop along the transmission line.	$\Delta V = IR$ $= 1200 \times 1.0$ $= 1200$ V
Determine the initial supply voltage.	$V_{\text{initial}} = V_{\text{supplied}} + \Delta V$ $= 250 \times 10^3 + 1200$ $= 251\,200$ V or 251.2 kV

Worked example: Try yourself 4.4.5

VOLTAGE DROP ALONG A TRANSMISSION LINE

Power is to be transmitted from the Loy Yang power station to Melbourne along a transmission line with a total resistance of 1.0 Ω. The current is 600 A. What voltage would be needed at the Loy Yang end of the transmission line to achieve a supply voltage of 500 kV?

LARGE-SCALE ELECTRICAL DISTRIBUTION SYSTEMS

Large-scale energy transmission is done through an interconnected grid between the power stations and the population centres where the bulk of the electrical energy is used. A wide-area synchronous grid, also known as an interconnection, directly connects a number of generators, delivering AC power with the same relative phase, to a large number of consumers.

No matter the source, the path the electrical power takes to the final consumer is very similar (see Figure 4.4.4). Step-up transformers in a large substation near the power station will raise the voltage from that initially generated to 240 000 V or 240 kV or more. The electrical power will then be carried via high-voltage transmission lines to a number of substations near key centres of demand. Substations with step-down transformers then reduce the voltage to more safe levels for distribution underground or via the standard 'electricity pole' you would be familiar with around city and country areas. Each group of 10–15 houses will be supplied by a smaller distribution transformer, mounted on the poles, which reduces the voltage down to the 240 V AC rms voltage that home and business installations are designed to run on (see Figure 4.4.4).

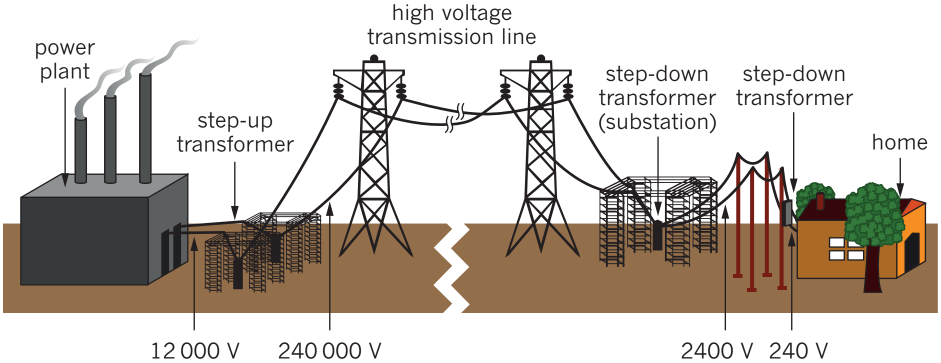


FIGURE 4.4.4 Transmitting electric power from generator to home uses AC power, so transformers can be used to minimise power losses through the system.

The use of AC as the standard for distribution allows highly efficient and relatively cheap transformers to convert the initial voltages created at the power station to much higher levels. The same power transmitted at a higher voltage requires less current and therefore less power loss. If it were not for this, the resistance of the transmission wires would need to be significantly reduced, which would require more copper in order to increase their cross-sectional area. This is both expensive and heavy. Less metal will make cables lighter and thinner, and the supporting towers themselves can be comparatively shorter, cheaper and lighter to build.

PHYSICS IN ACTION

The War of Currents

AC and DC power supplies have been in competition for nearly as long as humans have been generating electricity. The heated debates about the benefits and disadvantages of each type of current prompted what has been called the 'War of Currents' in the late 1800s. During this time Thomas Edison, an American inventor and businessmen, had created the Edison Electric Light Company that he hoped would supply electricity to large parts of America with his DC generators. Meanwhile, Nikola Tesla, a Serbian–American physicist, had invented the AC induction motor and, with financial support from George Westinghouse, hoped AC would become the dominant power supply. Ultimately, the ease with which AC could be stepped up using transformers for long-distance transmission with minimal power loss (as discussed in detail throughout this chapter) proved to be the prevailing benefit that led to AC winning the 'war'. However, in his attempt to win the competition, Edison attempted to portray the high voltage AC power as terrifyingly dangerous by using it to electrocute elephants and by inventing the AC-powered electric chair for the American government to execute prisoners on death row.

While AC power is now universal in large-scale power distributions, there is a limit to how high the voltage of an AC system can go and still be efficient. Above approximately 100 kV, corona loss (due to the high voltage ionising air molecules) begins to occur, and above 500 kV it no longer becomes feasible to transmit electric power due to these effects.

To transmit the same power as DC, an AC system would need to operate with a higher peak voltage. During the portion of the cycle when the AC is lower than peak voltages, efficiency is compromised because the higher the voltage the better. Up until recently the expense of alternative methods to raise and lower the voltage at either end of the transmission line more than negated this negative aspect of AC systems.

High DC voltage levels can now be reached more easily with new technology employing small, high-frequency switching converters. Projects such as the Three Gorges Dam in China (see Figure 4.4.5), and undersea transmission lines are now planning to use DC transmission. There are some other benefits, with many of the AC/DC transformers and three-phase industrial power currently in use becoming unnecessary. However, there is a whole range of other devices to be considered that would need to be allowed for, and major issues with safety. For example, safety switches won't work with DC power.

In some ways, the competition between Edison and Tesla continues. The Edison Electric Company merged in 1892 to become the General Electric Company, which exists to this day as one of the largest and most profitable companies in the world, while Westinghouse is still in business as a large home-appliance brand.

FIGURE 4.4.5 Transmitting projects at the Three Gorges Dam in China look to use DC transmission at higher voltages than is possible with AC to further reduce transmission losses.

4.4 Review

SUMMARY

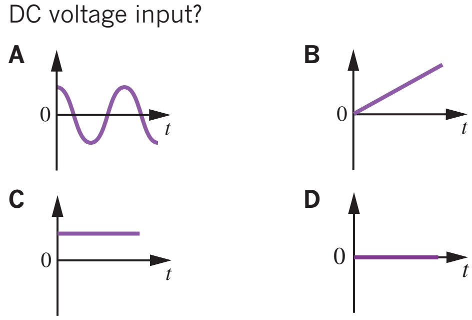
- A transformer works on the principle of a changing magnetic flux inducing an emf. No matter what the size or application, it will consist of two coils known as the primary and secondary coils.
- Ideal transformers are 100% efficient; real transformers are often over 99% efficient, and for this reason power losses within the transformer can be ignored in calculations.
- The transformer equation can be written in different versions but is based on:
  $\frac{V_2}{V_1} = \frac{N_2}{N_1}$
- A *step-up* transformer *increases* the secondary voltage compared with the primary voltage.
- A *step-down* transformer *decreases* the secondary voltage compared with the primary voltage.
- The transformer equation can also be written in terms of current, i.e.:
  $\frac{I_1}{I_2} = \frac{N_2}{N_1}$
- Transformers will not work with DC voltage since it has a constant, unchanging current that creates no change in magnetic flux.
- The power supplied in an electrical circuit is given by: $P = VI$
- The power lost in an electrical circuit is given by: $P = I^2R$
- The AC electrical supply from a generator is readily stepped up or down by transformers, hence AC is the preferred form of electrical energy in large-scale transmission systems.

KEY QUESTIONS

1 A non-ideal transformer has a slightly smaller power output from the secondary coil than input to the primary coil. The voltage and current in the primary coil are denoted V_1 and I_1 respectively. The voltage and current in the secondary coil are denoted V_2 and I_2 respectively. Which of the following expressions describes the power output in the secondary coil?

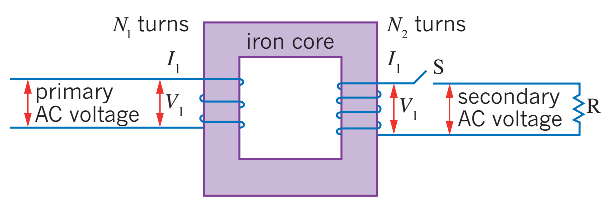
A V_1I_1
B V_2I_2
C V_1I_2
D $I_2{}^2R$

2 A voltage sensor is connected to the output of a transformer and a series of different inputs is used. Which of the following graphs is the most likely output displayed on a voltage graph for a steady DC voltage input?



3 A security light is operated from mains voltage 240 V rms through a step-down transformer with 800 turns on the primary winding. The security light operates normally on an rms voltage of 12 V.
How many turns are on the secondary coil?

4 The figure below depicts an iron core transformer. An alternating voltage applied to the primary coil produces a changing magnetic flux. The secondary circuit contains a switch, S, in series with a resistor, R. The number of turns in the primary coil is N_1 and in the secondary, N_2. The power in the first coil is P_1 and that in the second coil, P_2. Assume that this is an ideal transformer.



a Write an equation that defines the relationship between the power in the primary coil, P_1, and the power in the secondary coil, P_2.
b Write an equation that defines the relationship between the current in the secondary coil, I_2, and the current in the primary coil, I_1, in terms of the number of turns in each coil.

5 The primary windings of a transformer consist of 20 turns and the secondary of 200 turns. The primary rms voltage input is 8.0 V and a primary rms current of 2.0 A is flowing.
a What is the rms voltage across the load attached to the secondary coil?
b What power is being supplied to the load attached to the secondary coil?
c What is the rms current in the secondary coil?

4.4 Review *continued*

6 A security light is connected to a mains voltage of 240 V rms. It runs on a voltage of 12 V rms and an rms current of 2.0 A. A step-down transformer with 800 turns on the primary winding is used to reduce the voltage from the mains level to the required operating voltage. Assume that the light is operating normally and that there is no power loss in the transformer.

- **a** Calculate the number of turns, to the nearest whole number, in the secondary coil.
- **b** What is the value of the peak current in the primary coil?
- **c** Calculate the rms power input to the primary coil of the transformer.
- **d** During some routine maintenance work, the primary coil of the transformer for the security light is unplugged from the AC mains supply and plugged into a DC supply of 240 V instead. What is the new output (secondary) voltage?
  - **A** 0 V
  - **B** 12 V
  - **C** 24 V
  - **D** 240 V

7 A solar-powered generator produces 5.0 kW of electrical power at 500 V. This power is transmitted to a distant house via twin cables of total resistance 4.0 Ω. What is the total power loss in the cables?

8 A 100 km transmission line made from aluminium cable has a total resistance of 10 Ω. The line carries the electrical power from a 500 MW power station to a substation. If the line is operating at 250 kV, what is the power loss in the line?

9 A power station generates 500 MW of power to be used by a town 100 km away. The power lines between the power station and the town have a total resistance of 2.0 Ω.

- **a** If the power is transmitted at 100 kV, what current would be required?
- **b** What voltage would be available at the town? Give your answer in kilovolts (kV).

10 Power loss can be expressed by the formula $P = \frac{\Delta V^2}{R} = I^2R$. Therefore, select which of the following statements is true, and justify why the other response is incorrect:

- **A** The greater the voltage being transmitted in a transmission line, the greater the power loss.
- **B** The greater the current in the transmission line, the greater the power loss.

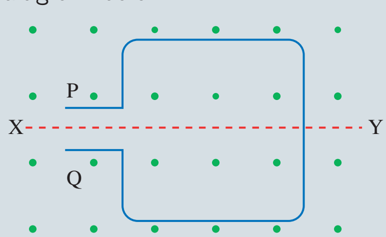
Chapter review

04

KEY TERMS

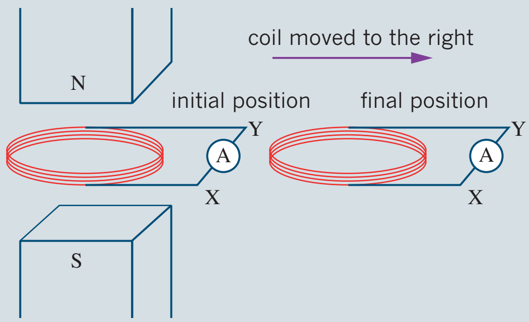
alternator	ideal transformer	slip rings
armature	induced current	split ring commutator
brushes	Lenz's law	step-down transformer
electromagnetic induction	magnetic flux	step-up transformer
Faraday's law	magnetic flux density	transformer
generator	root mean square	

1 A rectangular coil of area 40 cm^2 and resistance 1.0 Ω is located in a uniform magnetic field $B = 8.0 \times 10^{-4}$ T which is directed out of the page. The plane of the coil is initially perpendicular to the field as depicted in the diagram below.



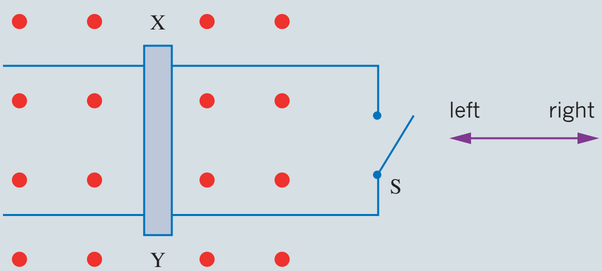
- **a** What is the magnitude of the emf induced in the coil when the strength of the magnetic field is doubled in a time of 1.0 ms?
- **b** What is the direction of the current caused by the induced emf in the coil when the strength of the magnetic field is doubled in a time of 1.0 ms?

2 During a physics experiment a student pulls a horizontal circular coil from between the poles of two magnets in 0.10 s. The initial position of the coil is entirely in the field while the final position is free of the field. The coil has 40 turns, each of radius 4.0 cm. The field strength between the magnets is 20 mT.



a What is the magnitude of the average emf induced in the coil as it is moved from its initial position to its final position?
b What is the direction of the current in the coil caused by the induced emf?

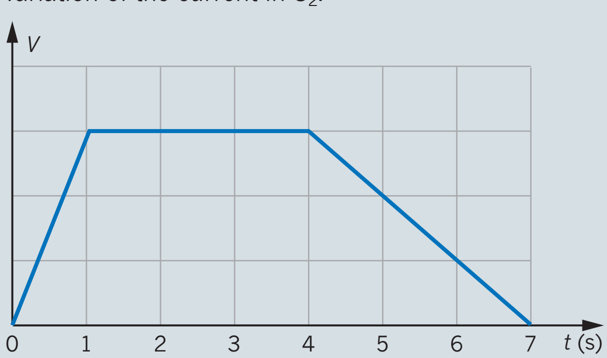
3 A copper rod, XY, of length 20 cm is free to move along a set of parallel conducting rails as shown in the following diagram. These rails are connected to a switch, S, which completes a circuit when it is closed. A uniform magnetic field of strength 10 mT, directed out of the page, is established perpendicular to the circuit. S is closed and the rod is moved to the right with a constant speed of 2.0 m s^{-1}.



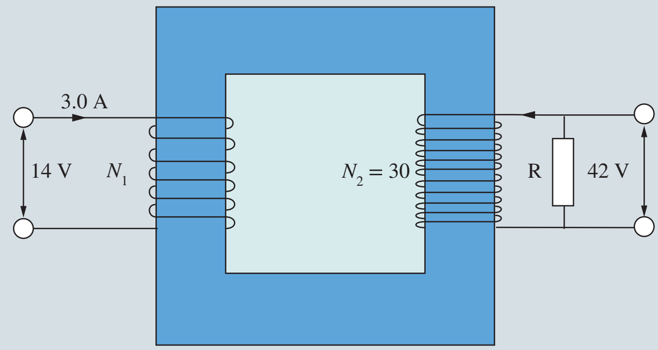
a What is the magnitude of the induced emf in the rod in mV?
b What is the direction of the current through the rod caused by the induced emf?

4 A ship with a vertical steel mast of length 8.0 m is travelling due west at 4.0 m s^{-1} in a region where the Earth's magnetic field is horizontal and is equal to 5.0×10^{-5} T north. What average emf, in mV, would be induced between the top and the bottom of the mast?

5 Coils S_1 and S_2 are close together and linked by a soft iron core. The emf in S_1 varies as shown in the graph below. Draw a line graph to show the shape of the variation of the current in S_2.



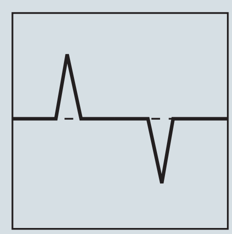
The following information relates to questions 6 and 7.
An ideal transformer is operating with an rms input voltage of 14 V and rms primary current of 3.0 A. The output voltage is 42 V. There are 30 turns in the secondary winding.



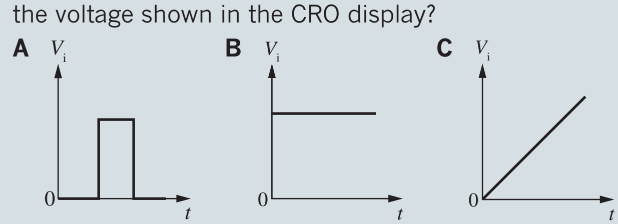
6 What is the rms output current?

7 How many turns are there in the primary coil?

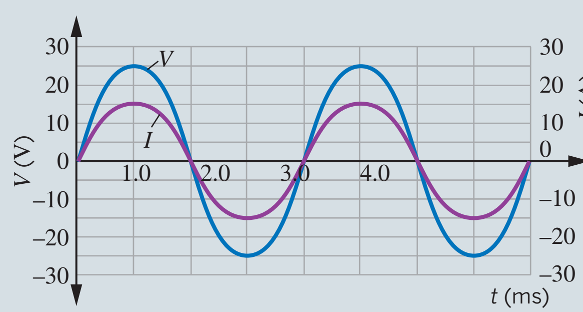
8 The following diagram shows a graph of induced voltage versus time as it appears on the screen of a CRO.



Which of the following input voltages would produce the voltage shown in the CRO display?



9 A physics student uses a voltage/current sensor to display the current, I, through, and the voltage, V, across, the output terminals of a small generator. The graph obtained from the display is shown below.



a What is the approximate rms voltage for the signal?
b Calculate the peak power output of the generator.

Chapter review *continued*

10 A student decides to test the power output of a new stereo amplifier. The maximum rms power output guaranteed by the manufacturer (assumed accurate) is 60 W. Which set of specifications is consistent with this power output?

peak–peak voltage (V)	peak–peak current (A)
A 20	3.0
B 40	6.0
C 40	12.0
D 20	6.0

The following information refers to questions 11 and 12.

A student builds a simple alternator consisting of a coil containing 500 turns, each of area 10 cm^2, mounted on an axis that can rotate between the poles of a permanent magnet of strength 80 mT. The alternator is rotated at a frequency of 50 Hz.

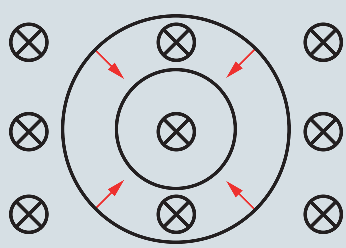
11 Find the average emf of the alternator.

12 Explain what the effect will be on the average emf when the frequency is doubled to 100 Hz.

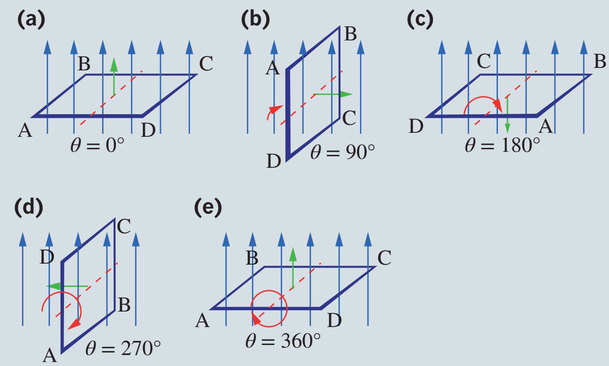
13 A generator is to be installed in a farm shed to provide 240 V power for the farmhouse. A twin-conductor power line with total resistance 8 Ω already exists between the shed and house. The farmer has seen a cheap 240 V DC generator advertised and is tempted to buy it.

Identify and explain two significant problems that you foresee with using the 240 V DC generator.

14 A coil in a magnetic field directed into the page is reduced in size. In what direction will the induced current flow in the coil while the coil is being reduced in size?



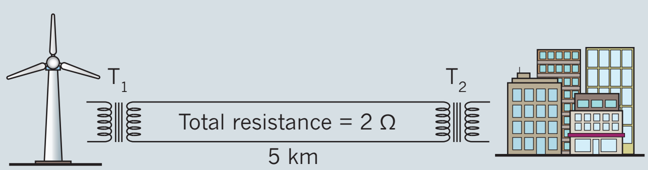
15 A single loop of wire is rotated within a magnetic field, *B*, as shown below.



While the coil is rotating, an emf will be generated as a result of *which* sides of the coil? Give a reason for your answer.

The following information relates to questions 16–19.

A wind turbine runs a 150 kW generator with an output voltage of 1000 V. The voltage is increased by a transformer T_1 to 10 000 V for transmission to a town 5 km away through power lines with a total resistance of 2 Ω. Another transformer, T_2, at the town reduces the voltage to 250 V. Assume that there is no power loss in the transformers (i.e. they are 'ideal').



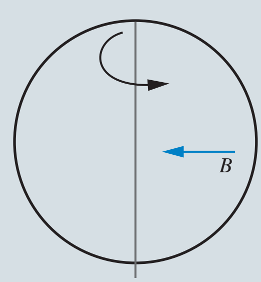
16 What is the current in the power lines?

17 What is the voltage at the input to the town transformer T_2?

18 How much power is lost in the power lines?

19 It is suggested that some money could be saved from the scheme by removing the first transformer. Explain, using appropriate calculations, whether this is a good plan.

20 A coil is rotated about its vertical axis such that the left-hand side would be coming out of the page and the right-hand side would be going into it. A magnetic field runs from right to left across the page. In what direction would the induced current in the coil flow?



UNIT 3 • Area of Study 2

REVIEW QUESTIONS

How are fields used to move electrical energy?

The following information relates to questions 1–4.

A rectangular loop of 100 turns is suspended in a magnetic field $B = 0.50$ T. The plane of the loop is parallel to the direction of the field. The dimensions of the loop are 20 cm perpendicular to the field lines and 10 cm parallel to them.

1 What amount of flux threads the loop in the position described above?

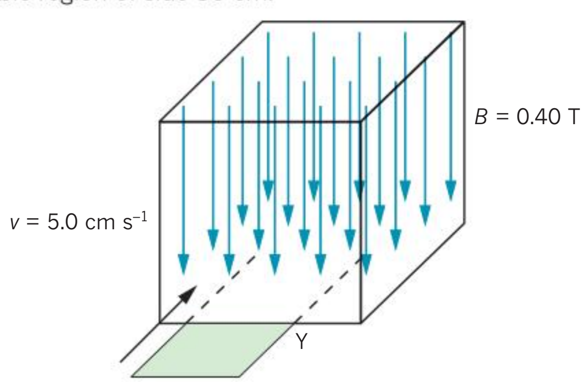
2 How can the amount of flux threading the loop be increased?

3 How should the plane of the loop and the magnetic field direction be arranged so the maximum possible flux threads the loop?

4 Calculate the maximum possible flux as described in Question 3.

The following information relates to questions 5–11.

A square conducting loop with sides 20 cm and resistance 0.50 Ω is moving with a constant horizontal velocity of 5.0 cm s^{-1} towards a region of uniform magnetic field of strength 0.40 T directed vertically downwards, as shown in the following diagram. The magnetic field is confined to a cubic region of side 30 cm.



5 Describe the direction of the induced current in the side XY of the loop just as it begins to enter the field. Justify your answer.

6 Calculate the average emf induced in the loop when it is halfway into the field.

7 What current flows in the loop when it is halfway into the field?

8 How much electrical power is consumed in the loop when it is halfway into the field?

9 What is the source of this power?

10 What is the average emf induced in the loop 5 s after it started to enter the cube? Justify your answer.

11 What is the direction of the induced current in the side XY just as it begins to emerge from the field? Justify your answer.

The following information relates to questions 12–16.

A rectangular conducting loop of dimensions 100 mm × 50 mm and resistance $R = 2.0$ Ω, is located with its plane perpendicular to a uniform magnetic field of strength $B = 1.0$ mT.

12 Calculate the magnitude of the magnetic flux Φ_B threading the loop.

13 The loop is rotated through an angle of 90° about an axis, so that its plane is now parallel to B. Determine the magnetic flux Φ_B threading the loop in the new position.

14 The time interval for the rotation $\Delta t = 2.0$ ms. Determine the average emf induced in the loop.

15 Determine the value of the average current induced in the loop during the rotation.

16 Will the current keep flowing once the rotation is complete and the loop is stationary? Explain your answer.

The following information relates to questions 17–18.

A 5.0 Ω coil, of 100 turns and radius 3.0 cm, is placed between the poles of a magnet so that the flux is a maximum through its area. The coil is connected to a sensitive current meter that has an internal resistance of 595 Ω. It is then moved out of the field of the magnet and it is found that an average current of 50 μA flows for 2 s.

17 Had the coil been moved out more quickly so that it was removed in only 0.5 s, what would have been the average current?

18 What is the strength of the magnetic field?

The following information relates to questions 19–21.

A physics student constructs a simple generator consisting of a coil of 400 turns. The coil is mounted on an axis perpendicular to a uniform magnetic field of strength $B = 50$ mT and rotated at a frequency $f = 100$ Hz. It is found that during the rotation, the peak voltage produced is 0.9 V.

19 Sketch a graph showing the voltage output of the generator for at least two full rotations of the coil. Include a scale on the time and voltage axes.

20 What is the RMS voltage generated?

21 The student now rotates the coil with a frequency $f = 200$ Hz. How would your answers to questions 19 and 20 be affected?

The following information relates to questions 22–26.

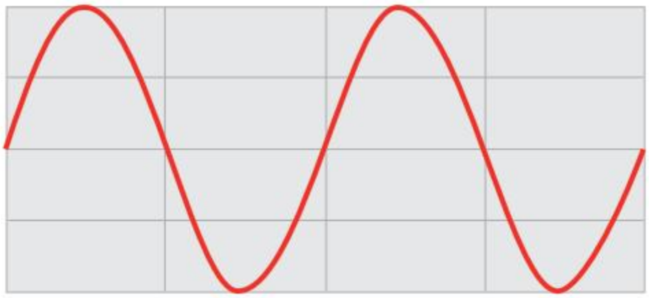
A generator is rotating at a rate of 3000 revolutions per minute. The magnetic field strength is 0.50 T. The total number of turns in the armature coils is $N = 200$, each of area $A = 100\ cm^2$.

22 Calculate the frequency of rotation of the generator.

23 Calculate the average emf generated during a quarter revolution of the generator coil.

The following graphs A–D and table apply to questions 24–26.

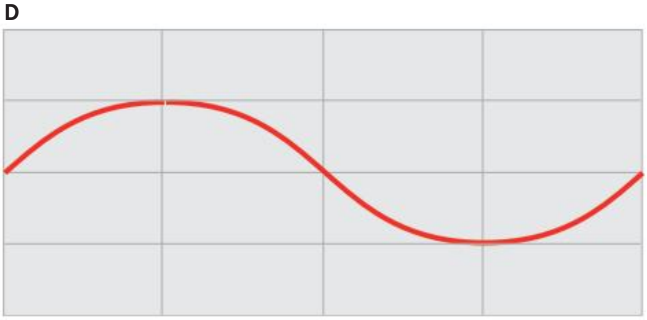
A



B

C

D



	f(Hz)	B(T)	N	A(cm²)
A	50	0.50	200	100
B	100	0.50	200	100
C	100	1.00	50	100
D	50	0.50	400	100

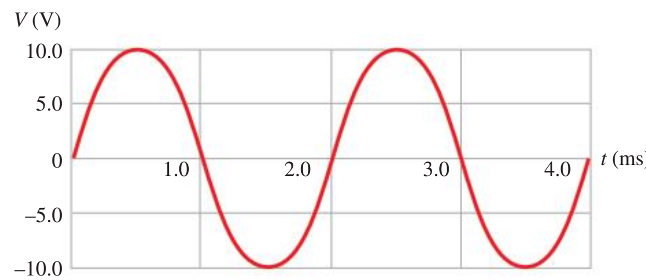
24 Which of the diagrams A–D best describes the display on the CRO when the generator is operating at a frequency of 100 Hz?

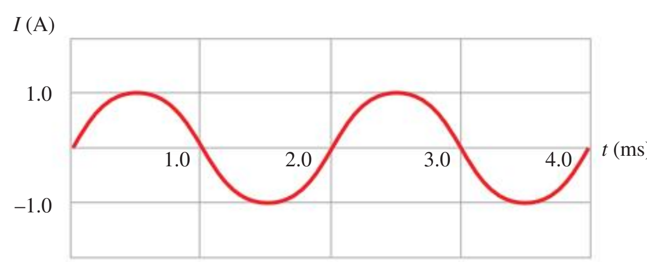
25 Which of the specifications in the table could produce a CRO display described by diagram A?

26 Which of the specifications in the table could produce a CRO display illustrated by diagram C?

The following information relates to questions 27–31.

The following diagram shows the voltage–time graph and corresponding current–time graph for an alternator that was built by a physics student as part of a research project.





	f(Hz)	B(T)	N	A(cm²)
A	50	0.50	200	100
B	100	0.50	200	100
C	100	1.00	50	100
D	50	0.50	400	100

27 What is the frequency of the voltage produced by the alternator?

28 What is the peak-to-peak output voltage of this alternator?

29 What is the rms output voltage of the alternator?

30 Calculate the rms output current of the alternator.

31 Calculate the rms output power of the alternator.

32 What feature distinguishes an alternator from a DC generator?

33 How does an alternator operate?

The following information relates to questions 34–38.

An ideal transformer is operating with peak input voltage of 600 V and an rms primary current of 2.0 A. The peak output voltage is 3000 V. There are 1000 turns in the secondary winding.

34 What is the rms output current?

35 What is the output peak-to-peak voltage?

36 How many turns are there in the primary winding?

37 Determine the RMS power consumed in the secondary circuit.

38 Calculate the peak power consumed in the secondary circuit.

39 Which of the following is the best description of how a transformer transfers electrical energy from the primary windings to the secondary windings?

A The current through the primary windings produces a constant electric field in the secondary windings.

B The current through the primary windings produces a steady magnetic field in the secondary windings.

C The current through the primary windings produces a changing magnetic field in the secondary windings.

40 When a transformer is plugged in to the 240 V mains but nothing is connected to the secondary coil, very little power is used. What is the best explanation for this?

A The primary and secondary coils are in series and so no current can flow in either if the secondary coil is open.

B There can be no magnetic flux generated in the transformer if the secondary coil has no current in it.

C The magnetic flux generated by the current in the primary produces an emf that opposes the applied voltage.

D The magnetic flux generated by the secondary coil almost balances out that due to the primary coil.

The following information relates to questions 41–50.

A farmer has installed a wind generator on a nearby hill, along with a power line consisting of two cables with a combined total resistance of 2.0 Ω. The output of the generator is given as 250 V AC (RMS) with a maximum power of 4000 W. She connects up the system and finds that the voltage at the house is indeed 250 V. However, when she turns on various appliances so that the generator is running at its maximum power output of 4000 W, she finds that the voltage supplied at the house is rather low.

41 Explain why the voltage dropped when the farmer turned on the appliances in the house.

42 Calculate the voltage and power at the house when the appliances are turned on.

She then decides to install ideal transformers at either end of the same power line so that the voltage transmitted from the generator end of the line in this system becomes 5000 V.

43 Describe the essential features of the types of transformers that are needed at either end of the power line.

For Questions 44–47, assume the generator is operating at full load, i.e. 4000 W.

44 What is the current in the power line now when the same appliances are turned on?

45 What is the voltage drop along the power line?

46 What is the power loss in the power line?

47 What voltage is delivered to the house?

48 What power is delivered to the house?

49 How do the power losses in the system without the transformers compare to the system with the transformers as a percentage of the power generated?

50 Explain why the system operated with much lower power losses when the voltage was transmitted at the higher voltage.

CHAPTER

05 Newtonian theories of motion

An understanding of forces and fields has allowed humans to land on the Moon and to explore the outer reaches of the solar system. Satellites in orbit around the Earth have changed the way people live.

These advances have been achieved using Newton's laws of motion, which were published in the 17th century. Newton suggested that it should be possible to put satellites in orbit around the Earth almost 300 years before it was technically possible to do so. While relativistic corrections introduced by Einstein are important in a limited number of contexts, Newton's description of gravitation and the laws governing motion are accurate enough for most practical purposes.

In this chapter Newton's laws will be used to analyse motion when two or more forces act on a body and how projectiles travel in the Earth's gravitational field. How forces keep objects travelling in a circular path will also be covered.

Key knowledge

By the end of this chapter, you will have covered material from the study of Newton's laws including how to:

- investigate and apply theoretically and practically Newton's three laws of motion in situations where two or more coplanar forces act along a straight line and in two dimensions
- investigate and analyse theoretically and practically the uniform circular motion of an object moving in a horizontal plane: $F_{net} = \frac{mv^2}{r}$, including:
  - a vehicle moving around a circular road
  - a vehicle moving around a banked track
  - an object on the end of a string
- model natural and artificial satellite motion as uniform circular motion
- investigate and apply theoretically Newton's second law to circular motion in a vertical plane (forces at the highest and lowest positions only)
- investigate and analyse theoretically and practically the motion of projectiles near Earth's surface, including a qualitative description of the effects of air resistance
- investigate and apply theoretically and practically the laws of energy and momentum conservation in isolated systems in one dimension.

5.1 Newton's laws of motion

On 14 July 2015, NASA's New Horizon's spacecraft (shown in Figure 5.1.1) sped past Pluto and sent back images to Earth that appeared on news broadcasts across the world. The trip had taken nine and a half years, and set the record for the fastest launch of a man-made object from Earth. The principles of physics on which this mission depended were published by Isaac Newton in 1687 in a set of laws that radically challenged the understanding of his time.

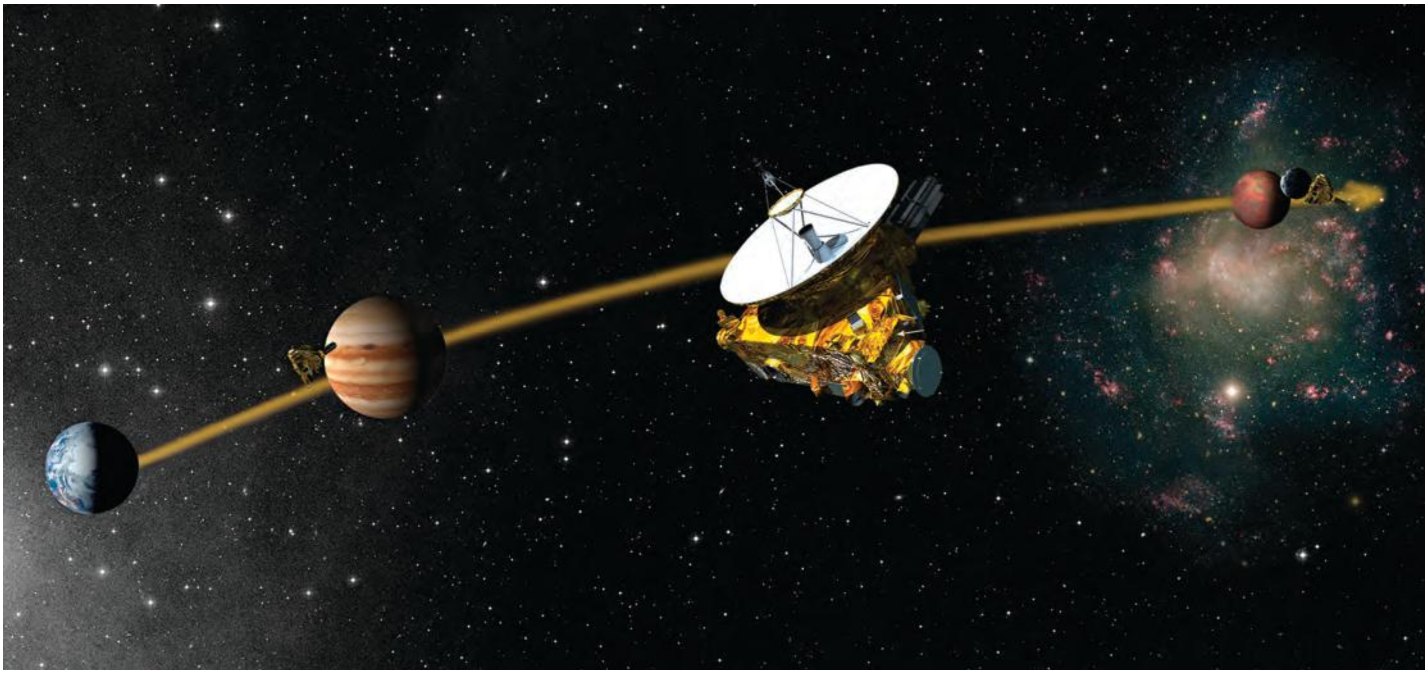


FIGURE 5.1.1 Artwork representing New Horizon's path to Pluto, flying past Jupiter on the way.

Newton's laws are, in fact, only an approximation and are superseded by Einstein's relativistic theories. In situations involving extremely high speeds (greater than 10% of the speed of light) or strong gravitational fields, Newton's laws become imprecise, and Einstein's theories must be used instead. However, Newton's laws are not obsolete. For the most part, Newton's laws remain invaluable for describing the motion of objects as diverse as planets and ping-pong balls.

NEWTON'S THREE LAWS OF MOTION

Newton's laws describe how forces can be used to explain the motion of bodies. The first law describes what happens to a body when there is no net force on it. The second law explains motion when there is an unbalanced force acting, and the third states that all forces act in action–reaction pairs.

Newton's first law

Newton's first law states that every object continues to be at rest, or continues with constant velocity, unless it experiences an unbalanced force. This is also called the law of inertia. An object that is moving at constant velocity will keep moving. This is seldom observed in everyday life due to the presence of forces such as friction and air resistance which will eventually slow the motion of the object. To maintain constant motion, frictional forces must be balanced with some other force. For example, an object can keep moving at a constant velocity if it is driven by a motor.

An object that is stationary will remain stationary while the forces acting on it are balanced. For example, an object will fall due to the force of gravity, but will remain at rest when this force is balanced by the normal reaction force applied by a table.

Newton's second law

Newton's second law states that the acceleration of a body experiencing an unbalanced force is directly proportional to the net force acting on it and inversely proportional to the mass of the body:

$F_{net} = ma$

In other words, an object will accelerate at a greater rate when the force acting on it is increased; bigger forces mean greater acceleration. Heavy objects are harder to accelerate than lighter ones. So the rate of acceleration decreases for larger masses.

The net or resultant force, F_{net}, is measured in newtons (N), when the mass is measured in kilograms (kg) and the acceleration, a, is measured in metres per second squared (m s^{-2}).

Newton's third law

Newton's third law states that when one body exerts a force on another body (an action force), the second body exerts an equal force in the opposite direction on the first (the reaction force):

$F_{\text{on A by B}} = -F_{\text{on B by A}}$

To simplify the notation, this text will use the convention

$F_{AB} = F_{\text{on A by B}}$

Hence the first subscript always shows the body experiencing the force.

It is important to note that action–reaction pairs can never be added together, because they act on different bodies (see Figure 5.1.2). The forces in an action–reaction pair:

- are the same **magnitude** (size)
- act in opposite directions and
- are exerted on two different objects.

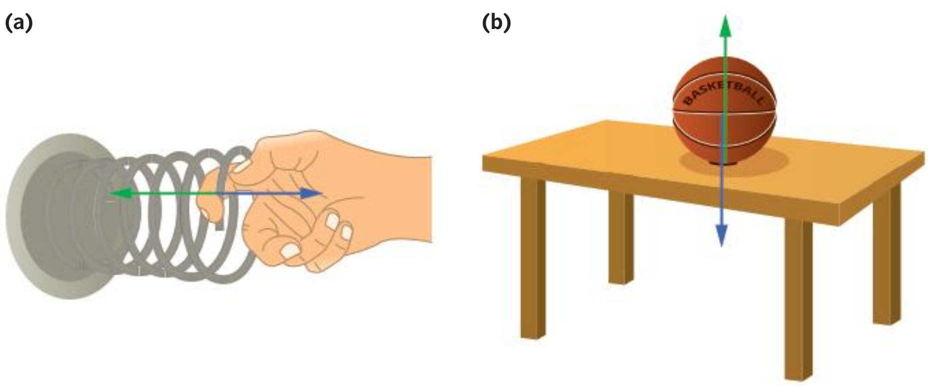


FIGURE 5.1.2 Figure (a) shows an action–reaction pair. The hand pulls on the spring and the spring pulls back on the hand with an equal and opposite force. Figure (b) does not show an action–reaction pair. This is because the force due to gravity and the normal reaction force both act on the same object, the basketball.

While the force is the same size on both objects, the resulting acceleration may not be. That's because the rate of acceleration depends on the mass of the objects concerned (from Newton's second law). Sometimes, when the objects have very different masses, the effect of one force in the pair is much more noticeable. For example, if you stub your toe on a large, heavy rock, the force exerted on your toe by the rock causes your foot to decelerate significantly. The equal and opposite force exerted by your toe on the rock, does not cause significant acceleration of the rock, because of its much greater mass.

PHYSICSFILE

Tethered spacewalks

When stationed on the International Space Station (ISS), astronauts are often required to conduct spacewalks—that is, they need to complete tasks outside of their spacecraft. During spacewalks, astronauts are tethered (attached) to their vehicles. If they weren't they would float off into space (remember Newton's first law of motion!). All of the astronaut's tools are attached to their spacesuits, otherwise, they too would float off into space. If an astronaut were to become accidentally untethered, it could be a disaster. Without a surface to push against, the astronaut would float off into space without the ability to return to the spacecraft. As a safety precaution, every astronaut is fitted with a small jet pack that they can fire to propel and manoeuvre them back to their vehicle. The jet pack propels the astronaut forward when it is fired backwards (remember Newton's third law of motion!).

Worked example 5.1.1

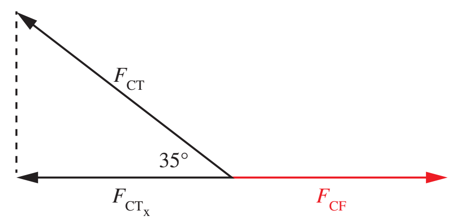
APPLICATION OF NEWTON'S FIRST AND THIRD LAWS

A toddler drags his 4.5 kg cart of blocks across a floor at a constant speed of 0.75 m s^{-1}. The force of friction between the cart and the floor is 5.0 N, and it is being pulled by a handle which is at an angle of 35° above the horizontal.

a Calculate the net force on the cart.

Thinking	Working
The cart has constant velocity. According to Newton's first law, the net force acting on the cart must be zero.	$F_{c\ net} = 0$ N

b Calculate the force that the toddler exerts on the cart.

Thinking	Working
Draw a diagram.	
If the net force is zero then the horizontal forces must be in balance. Therefore, the horizontal component of the force on the cart by the toddler, F_{CTx}, is equal to the magnitude of the frictional force, F_{CF}.	$F_{CTx} = F_{CT} \cos 35° = F_{CF}$ $F_{CT} \cos 35° = 5.0$ N $F_{CT} = \frac{5.0}{\cos 35°} = 6.1$ N

c Determine the force that the cart exerts on the toddler.

Thinking	Working
Apply Newton's third law to find the force on the toddler by the cart.	According to Newton's third law, the force on the cart by the toddler is equal and opposite to the force on the toddler by the cart. $F_{CT} = -F_{TC}$ Since the force on the cart is at an angle of 35° above the horizontal, the force of the cart on the toddler is 6.1 N at an angle of 35° below the horizontal.

Worked example: Try yourself 5.1.1

APPLICATION OF NEWTON'S FIRST AND THIRD LAWS

The toddler adds extra blocks to the cart and drags it across the floor more slowly. The 5.5 kg cart travels at a constant speed of 0.65 m s^{-1}. The force of friction between the cart and the floor is 5.2 N and the handle is now at an angle of 30° above the horizontal.

a Calculate the net force on the cart.

b Calculate the force that the toddler exerts on the cart.

c Calculate the force that the cart exerts on the toddler.

Applying Newton's first or second laws

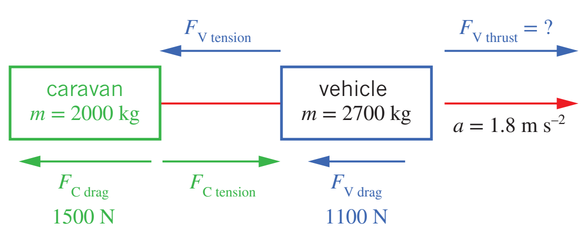
When solving motion problems, a key strategy is to determine whether Newton's first or second law should be applied. In the following examples the objects in the questions are accelerating, hence the second law should be used, and the net force is proportional to the acceleration. In problems involving connected bodies, both the whole system and each component of the system have the same acceleration.

Worked example 5.1.2

APPLICATION OF NEWTON'S LAWS

A vehicle towing a caravan accelerates at 1.8 m s^{-2} in order to overtake a car in front. The vehicle's mass is 2700 kg and the caravan's mass is 2000 kg. The drag forces on the vehicle are 1100 N, while the drag forces on the caravan are 1500 N.

a Calculate the thrust of the engine.

Thinking	Working
Draw a sketch showing all forces acting.	
Since there is an acceleration, Newton's second law may be applied to the whole system. Note that the caravan and vehicle are joined by the coupling and so the tension forces are not included at this stage. Consider the system as a whole.	$F_{system} = m_{system}a$ $F_{V\ thrust} - F_{V\ drag} - F_{C\ drag} = (m_C + m_V)\ a$ $F_{V\ thrust} - 1100 - 1500 = (2000 + 2700) \times 1.8$ $F_{V\ thrust} = 1.1 \times 10^4$ N in the direction of motion

b Calculate the magnitude of the tension in the coupling.

Thinking	Working
Consider only one part of the system, for example the caravan, once again applying Newton's second law.	$F_{C\ net} = m_C a$ $F_{C\ tension} - F_{C\ drag} = m_C a$ $F_{C\ tension} = 2000 \times 1.8 + 1500$ $= 5.1 \times 10^3$ N

Worked example: Try yourself 5.1.2

APPLICATION OF NEWTON'S LAWS

A vehicle towing a trailer accelerates at 2.8 m s^{-2} in order to overtake a car in front. The vehicle's mass is 2700 kg and the trailer's mass is 600 kg. The drag forces on the vehicle are 1100 N, and the drag forces on the caravan are 500 N.

a Calculate the thrust of the engine.

b Calculate the magnitude of the tension in the coupling.

THE NORMAL FORCE

One reaction force deserves a special mention. When an object exerts a force on a surface, the surface exerts a reaction force on the object that is normal (at right angles) to the surface.

For example, the block in Figure 5.1.3(a) exerts a force on the surface because it is attracted towards the centre of the Earth by gravity. The surface exerts a **normal reaction force** on the block. The weight F_g is thus balanced by F_N as shown in the figure. There is no net force on the block, and so Newton's first law applies and the object remains stationary.

On an **inclined plane**, F_N is at an angle to F_g. There is a net force down the slope and the block accelerates as predicted by Newton's second law.

Another way of viewing the forces along the inclined plane is to resolve the weight vector into two components: one perpendicular (at right angles) to the slope, and one parallel to the slope as shown in Figure 5.1.4. The component perpendicular to the surface is balanced by the normal force F_N. The component of the weight directed along the slope is the force that actually causes the acceleration.

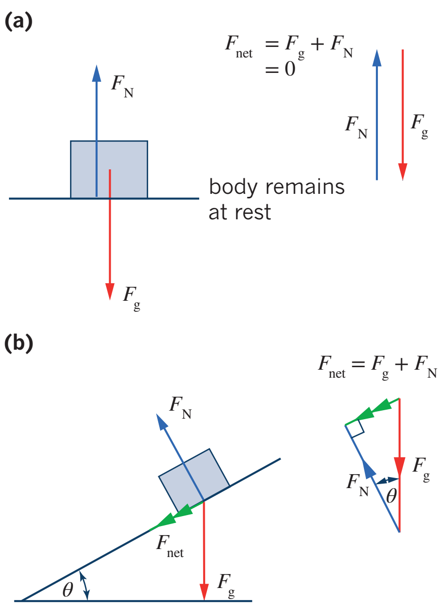


FIGURE 5.1.3 (a) Block on level surface: the net force is zero as F_N and F_g cancel. (b) Block on incline: $F_N = F_g \cos \theta$, and the net force is given by $F_{net} = F_g + F_N$ added as vectors.

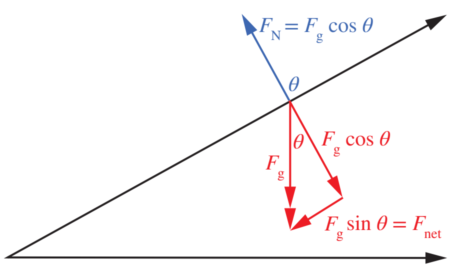
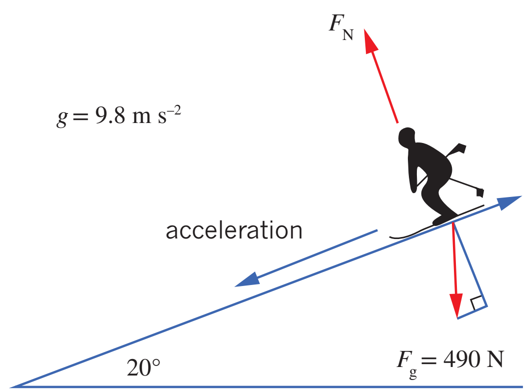


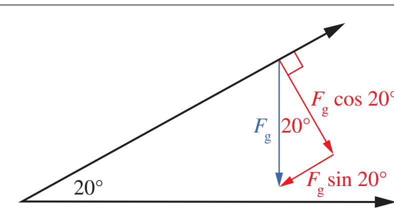
FIGURE 5.1.4 Block on an incline: the weight force can be resolved into a force perpendicular to the surface and a force parallel to the surface.

Worked example 5.1.3

INCLINED PLANES

A skier of mass 50 kg is skiing down an icy slope that is inclined at 20° to the horizontal. Assume that friction is negligible and that the acceleration due to gravity is 9.8 m s^{-2}.



a Determine the components of the weight of the skier perpendicular to the slope and parallel to the slope.	
Thinking	**Working**
Draw a sketch including the values provided.	
Resolve the weight into a component perpendicular to the slope.	The perpendicular component is: $F_{\perp} = F_g \cos 20°$ $= 490 \cos 20°$ $= 460$ N
Resolve the weight into a component parallel to the slope.	The parallel component is $F = F_g \sin 20°$ $= 490 \sin 20°$ $= 168$ N

b Determine the normal force that acts on the skier.	
Thinking	**Working**
The normal force is equal in magnitude to the perpendicular component of the weight force.	$F_N = 460$ N

c Calculate the acceleration of the skier down the slope.	
Thinking	**Working**
Apply Newton's second law. The net force along the plane is the component of the weight parallel to the slope.	$a = \frac{F_{net}}{m}$ $= \frac{168}{50}$ $= 3.36$ m s^{-2} down the slope

Worked example: Try yourself 5.1.3

INCLINED PLANES

A much heavier skier of mass 85 kg travels down the same icy slope inclined at 20° to the horizontal. Assume that friction is negligible and that the acceleration due to gravity is 9.8 m s^{-2}.

a Determine the components of the weight of the skier perpendicular to the slope and parallel to the slope.

b Determine the normal force that acts on the skier.

c Calculate the acceleration of the skier down the slope.

Aside from rounding differences, the acceleration calculated in the Worked example and Try yourself questions above were equal. That's because acceleration is independent of the mass of the object. This is because:

$$a = \frac{F_{net}}{m} = \frac{mg \sin \theta}{m} = g \sin \theta$$

STRATEGIES FOR SOLVING FORCE AND MOTION PROBLEMS

Where forces on a body are given, Newton's laws can be applied.

Two questions can be asked:

1 Is the object described as stationary or travelling at constant velocity? In this case $F_{net} = 0$.

2 Is the object accelerating? In this case, $F_{net} = ma$.

When dealing with connected bodies, consider the whole system first, and then consider the separate parts of the system.

For coplanar forces that are not aligned, resolve forces into components.

Newton's second law can be used to find the acceleration of an object, after which the other equations of motion may be used to find quantities such as displacement and final velocity.

5.1 Review

SUMMARY

- Newton's first law states that every object continues to be at rest, or continues with constant velocity, unless it experiences an unbalanced force. This is also called the law of inertia.
- Newton's second law states that the acceleration of a body experiencing an unbalanced force is directly proportional to the net force and inversely proportional to the mass of the body: $F_{net} = ma$.
- Newton's third law states that when one body exerts a force on another body (an action force), the second body exerts an equal force in the opposite direction on the first (the reaction force): $F_{AB} = -F_{BA}$.
- The forces in an action–reaction pair are the same magnitude, act in opposite directions and are exerted on two different objects.
- A normal reaction force, F_N, acts between an object and a surface, at right angles to the surface.
  - On a horizontal surface, $F_N = F_g$ and the object is stationary.
  - On an inclined surface, F_N is equal and opposite to the component of the weight force acting perpendicular to the plane: $F_N = F_g \cos \theta$.
- The net force acting on an object on a plane inclined at an angle θ is: $F_{net} = F_g \sin \theta$ when friction is negligible.

KEY QUESTIONS

1 Phil is standing inside a tram when it starts off suddenly. Len, who was sitting down, commented that Phil was 'thrown backwards' as the tram started moving. Is this a correct statement? Explain in terms of Newton's laws.

2 A table-tennis ball of mass 10 g is falling towards the ground with a constant speed of 8.2 m s^{-1}. Calculate the magnitude and direction of the air resistance force acting on the ball.

3 Ishtar is riding a motorised scooter along a level bike path. The combined mass of Ishtar and her scooter is 80 kg. The frictional and drag forces that are acting total to 45 N. What is the magnitude of the driving force being provided by the motor if she is:

a moving with a constant speed of 10 m s^{-1}

b accelerating at 1.5 m s^{-2}?

4 A cyclist and his bike have a combined mass of 80 kg. When starting off from traffic lights, the cyclist accelerates uniformly and reaches a speed of 7.5 m s^{-1} in 5.0 s.

a What is the acceleration during this time?

b Calculate the driving force being provided by the cyclist's legs as he starts off. Assume that drag forces are negligible during this time.

c The cyclist now rides along with a constant speed of 15 m s^{-1}. Assuming that the force being provided by his legs is now 60 N, determine the magnitude of the drag forces that are acting.

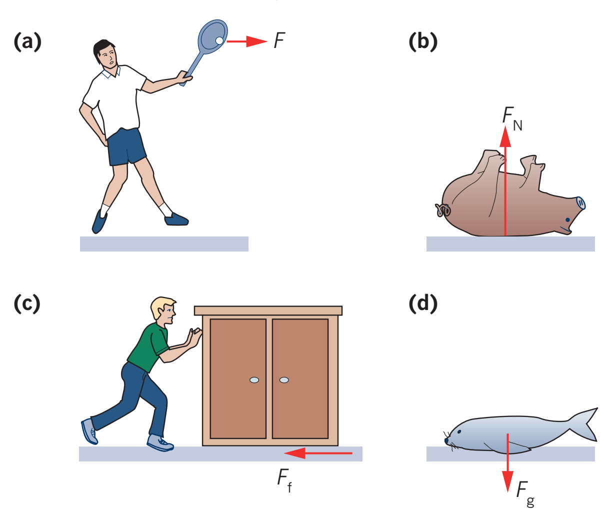
5 During preseason football training, Matt was required to run with a bag of sand dragging behind him. The bag of mass 50 kg was attached to a rope, which made an angle of 25° to the horizontal. When Matt ran with a constant speed of 4.0 m s^{-1}, a frictional force of 60 N was acting on the bag.

a What was the net force acting on the bag of sand?

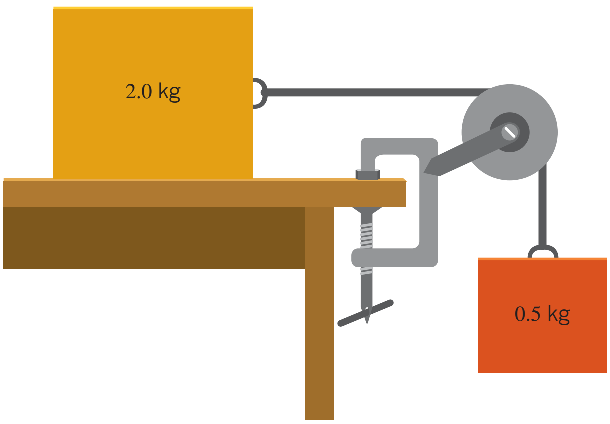
b Calculate the size of the tension force acting in the rope.

c What was the magnitude of the force that the rope exerted on Matt as he ran?

6 Complete each of these force diagrams, showing the reaction pair to the action force that is shown. For each force that you draw, state what the force is acting on and what is providing the force.



7 A block on the table is accelerated by a falling weight as shown in the figure below. Calculate the tension in the cord if the block experiences a frictional force of 1.5 N as it slides on the table.

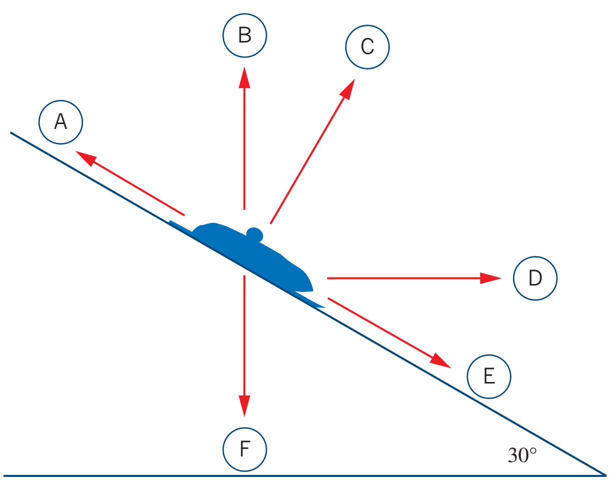


8 A 2000 kg tractor tows a 250 kg tree stump along a level surface with an acceleration of 1.5 m s^{-2}. The frictional drag on the tractor is 600 N and the drag on the tree is 1025 N.

a Calculate the thrust required by the tractor engine.

b The breaking strength of the towrope is 1500 N. Complete a calculation to determine whether or not the rope will break.

9 Kirsty is riding in a bobsled that is sliding down a snow-covered hill with a slope of 30° to the horizontal. The total mass of the sled and Kirsty is 100 kg. Initially the brakes are on and the sled moves down the hill with a constant velocity.



a Which one of the arrows (A–F) best represents the direction of the frictional force acting on the sled?

b Which one of the arrows (A–F) best represents the direction of the normal force acting on the sled?

c Calculate the net frictional force acting on the sled.

d Kirsty then releases the brakes and the sled accelerates. What is the magnitude of her initial acceleration?

e Finally, Kirsty rides the bobsled down the same slope but with the brakes off, so friction can be ignored. It now has an extra passenger so that its total mass is now 140 kg. How will this affect the acceleration of the bobsled?

10 Which of the following statements describe the forces acting on an object on a stationary level surface? (More than one correct answer is possible.)

A The normal force is always perpendicular to the surface.

B The normal force is always equal in magnitude to the weight.

C The normal force and the weight are action–reaction pairs.

D The normal force and the weight cancel out.

11 Which of the following statements describes the forces acting on an object on a plane inclined at an angle θ?

A The normal force is always perpendicular to the surface.

B The normal force is equal in magnitude to the weight.

C The normal force and the weight cancel out.

D In the absence of friction, a component of the normal force causes the object to accelerate down the slope.

5.2 Circular motion in a horizontal plane

Circular motion is common throughout the universe. On a small scale, this could involve children moving in a circular path on a fair ride (see Figure 5.2.1) or passengers in a car as it travels around a roundabout. In athletics, hammer throwers swing the hammer in a circular path before releasing it at high speed. On a much larger scale, the planets orbit the Sun in roughly circular paths; and on an even grander scale, stars can travel in circular paths around the centres of their galaxies. This section explains the nature of circular motion in a horizontal plane, and applies Newton's first and second laws to different circular-motion problems.

FIGURE 5.2.1 The people on this ride are travelling in a circular path at high speed.

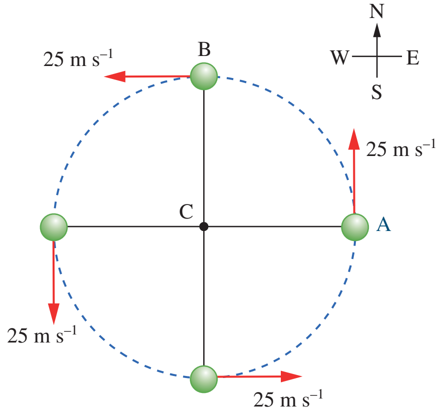


FIGURE 5.2.2 The velocity of the hammer (steel ball) at any instant is tangential to its path and is continually changing even though it has constant speed. This changing velocity means that the hammer is accelerating.

UNIFORM CIRCULAR MOTION

Figure 5.2.2 shows an athlete in a hammer throw event, swinging a steel ball in a horizontal circle with a constant speed of 25 m s^{-1}. As the hammer travels in its circular path, its *speed is constant,* but its *velocity is continually changing.*

Remember that velocity is a vector. Since the direction of the hammer is changing, so too is its velocity, even though its speed is not changing.

The velocity of the hammer at any instant is **tangential** (at a tangent) to its path. At one instant, the hammer is travelling at 25 m s^{-1} north, then an instant later at 25 m s^{-1} west, then 25 m s^{-1} south, and so on.

PERIOD AND FREQUENCY

Imagine that an object is moving in a circular path with a constant speed, v, and a radius of r metres, and it takes T seconds to complete one revolution. The time required to travel once around the circle is called the **period**, T, of the motion. The number of rotations each second is the **frequency**, f.

$f = \frac{1}{T}$ and $T = \frac{1}{f}$

where f is the frequency (Hz)

T is the period (s)

SPEED

An object that travels in a circle will travel a distance equal to the circumference of the circle, $C = 2\pi r$, with each revolution (see Figure 5.2.3). Given that the time for each revolution is the period, T, the average speed of the object is:

$$\text{speed} = \frac{\text{distance}}{\text{time}} = \frac{\text{circumference}}{\text{period}}$$

In circular motion, this equation is represented as follows.

The average speed of an object moving in a circular path is:

$$v = \frac{2\pi r}{T}$$

where v is the speed (m s^{-1})

r is the radius of the circle (m)

T is the period of motion (s)

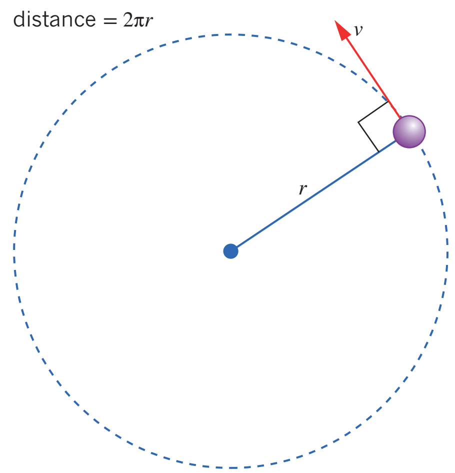


FIGURE 5.2.3 The average speed of an object moving in a circular path is given by the distance travelled in one revolution (the circumference) divided by the time taken (the period, T).

PHYSICSFILE

Wind generators

The wind generator in Figure 5.2.4 is part of a wind farm at Macarthur in south-west Victoria. This is the largest wind farm in Australia and consists of 140 turbines. The towers are 85 m high. Each blade is 55 m long and they rotate at a maximum rate of 20 revolutions per minute. From this information, you should be able to calculate that the tip of each blade is travelling at around 400 km h^{-1}!

FIGURE 5.2.4 The tips of these wind-generator blades are travelling in circular paths at speeds of around 400 km h^{-1}.

Worked example 5.2.1

CALCULATING SPEED

A wind turbine has blades 55.0 m in length that rotate at a frequency of 20 revolutions per minute. At what speed do the tips of the blades travel? Express your answer in km h^{-1}.	
Thinking	**Working**
Calculate the period, T. Remember to express frequency in the correct units. Alternatively, recognise that 20 revolutions in 60 s means that each revolution takes 3 s.	20 revolutions per minute = $\frac{20}{60}$ = 0.333 Hz $T = \frac{1}{f}$ $= \frac{1}{0.333} = 3$ s
Substitute r and T into the formula for speed and solve for v.	$v = \frac{2\pi r}{T}$ $= \frac{2 \times \pi \times 55.0}{3}$ $= 115.2$ m s^{-1}
Convert m s^{-1} into km h^{-1} by multiplying by 3.6.	$115.2 \times 3.6 = 415$ km h^{-1}

Worked Example: Try yourself 5.2.1

CALCULATING SPEED

A water wheel has blades 2.0 m in length that rotate at a frequency of 10 revolutions per minute. At what speed do the tips of the blades travel? Express your answer in km h^{-1}.

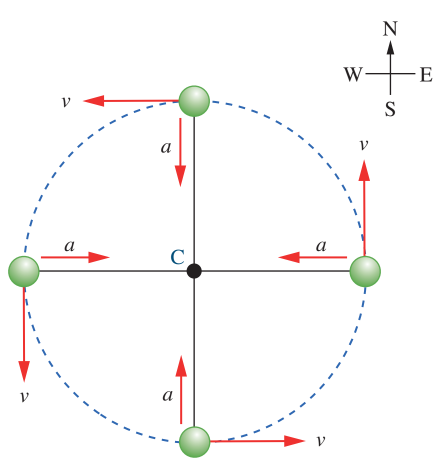


FIGURE 5.2.5 A body moving in a circular path has an acceleration towards the centre of the circle. This is known as a centripetal acceleration.

CENTRIPETAL ACCELERATION

When objects travel in circular paths, they can have a constant speed, yet at the same time have a velocity that is changing. This seeming contradiction arises because speed is a scalar quantity whereas velocity is a vector.

Since the velocity of the object is changing, it is accelerating even though its speed is not changing. The object is continually deviating inwards from its straight-line direction and so has an acceleration towards the centre. This acceleration is known as **centripetal acceleration**, a. In Figure 5.2.5, the velocity vector of an object travelling in a circular path is shown with an arrow labelled v. Notice how it is at a tangent to the circular path. The acceleration, a, is towards the centre of the circular path.

However, as Figure 5.2.5 shows, even though the object is accelerating towards the centre of the circle, it never gets any closer to the centre. This is the same principle that applies to satellites in orbit, which were studied in Chapter 3.

The centripetal acceleration, a, of an object moving in a circular path of radius r with a velocity v can be found from the relationship:

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

A substitution can be made for the speed of the object in this equation.

$$v = \frac{2\pi r}{T}$$

so

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

$$= \left(\frac{2\pi r}{T}\right)^2 \times \frac{1}{r}$$

$$= \frac{4\pi^2 r}{T^2}$$

Centripetal acceleration is always directed towards the centre of the circular path and is given by:

$$a = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2}$$

where a is the centripetal acceleration (m s^{-2})

v is the speed (m s^{-1})

r is the radius of the circle (m)

T is the period of motion (s)

FORCES THAT CAUSE CIRCULAR MOTION

As with all forms of motion, an analysis of the forces that are acting is needed to understand why circular motion occurs. In the hammer-throw event described earlier in this section, the hammer ball is continually accelerating. It follows from Newton's second law that there must be a net unbalanced force continuously acting on it. The net unbalanced force that gives the hammer ball its acceleration towards the centre of the circle is known as a **centripetal force**.

In every case of circular motion, a real force is necessary to provide the centripetal force. The force acts in the same direction as the acceleration, that is, towards the centre of the circle. This centripetal force can be provided in a number of ways. For the hammer in Figure 5.2.6, the centripetal force is the tension force in the cable. Other examples of centripetal force are also shown in Figure 5.2.6.

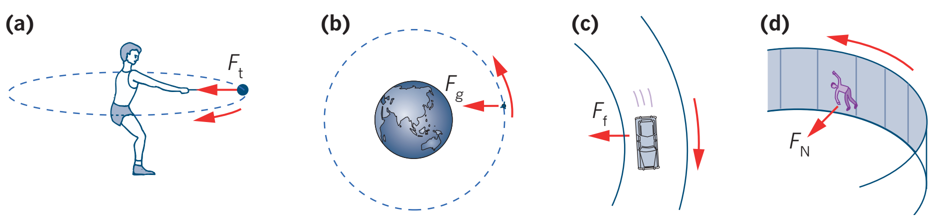


FIGURE 5.2.6 (a) In a hammer throw, tension in the cable provides the centripetal force. (b) For planets and satellites, the gravitational attraction to the central body provides the centripetal force. (c) For a car on a curved road, the friction between the tyres and the road provides the centripetal force. (d) For a person in the Gravitron ride, it is the normal force from the wall that provides the centripetal force.

Now, consider the consequences if the unbalanced force ceases to act. In the example of the hammer thrower, if the tension in the wire became zero because the thrower released the ball, there is no longer a force causing the ball to change direction. The result is that the ball then moves in a straight line tangential to its circular path, as would be expected from Newton's first law.

Centripetal force is given by:

$$F_{net} = ma = \frac{mv^2}{r} = \frac{4\pi^2 rm}{T^2}$$

where F_{net} is the net or centripetal force on the object (N)

m is the mass (kg)

a is the acceleration (m s^{-2})

v is the speed (m s^{-1})

r is the radius of the circle (m)

T is the period of motion (s)

Worked example 5.2.2

CENTRIPETAL FORCES

An athlete in a hammer throw event is swinging the ball of mass 7.0 kg in a horizontal circular path. The ball is moving at 20 m s^{-1} in a circle of radius 1.6 m.

a Calculate the magnitude of the acceleration of the ball.	
Thinking	**Working**
As the object is moving in a circular path, the centripetal acceleration is towards the centre of the circle. To find the magnitude of this acceleration, write down the other variables that are given.	$v = 20$ m s^{-1} $r = 1.6$ m $a = ?$
Find the equation for centripetal acceleration that fits the information you have, and substitute the values.	$a = \frac{v^2}{r}$ $= \frac{20^2}{1.6}$ $= 250$ m s^{-2}
Calculate the magnitude only, so no direction is needed in the answer.	The acceleration of the ball is 250 m s^{-2}.

b Calculate the magnitude of the tensile (tension) force acting in the wire.	
Thinking	**Working**
Identify the unbalanced force that is causing the object to move in a circular path. Write down the information that you are given.	$m = 7.0$ kg $a = 250$ m s^{-2} $F_{net} = ?$
Select the equation for centripetal force, and substitute the variables you have.	Equation for centripetal forces: $F_{net} = ma$ $= 7.0 \times 250$ $= 1.8 \times 10^3$ N
Calculate the magnitude only, so no direction is needed in the answer.	The force of tension in the wire is the unbalanced force that is causing the ball to accelerate. Tensile force $F_T = 1.8 \times 10^3$ N

Worked example: Try yourself 5.2.2

CENTRIPETAL FORCES

An athlete in a hammer throw event is swinging the ball of mass 7.0 kg in a horizontal circular path. The ball is moving at 25 m s^{-1} in a circle of radius 1.2 m.

a Calculate the magnitude of the acceleration of the ball.

b Calculate the magnitude of the tensile force acting in the wire.

PHYSICS IN ACTION

The Gravitron

When a car turns sharply to the left, the passengers in the car seem to sway to the right inside the car. Many mistakenly think that a force to the right is acting. In fact, the passengers are simply maintaining their motion in the original direction of the car as described by Newton's first law, that is, they are experiencing inertia. If the passengers are (unwisely) not wearing seatbelts, they will be squashed against the right-hand door as the car turns. This will exert a large force to the left on them, which causes them to move to the left.

People moving rapidly in circular paths also mistakenly think that there is an outward force acting on them. For example, riders on the Gravitron (also known as the Vortex or Rotor), like those in Figure 5.2.7, will 'feel' a force pushing them into the wall. This outwards force is commonly known as a centrifugal (meaning 'centre-fleeing') force. This force does not actually exist in an inertial frame of reference. The riders think that it does this because they are in the rotating frame of reference. From outside the Gravitron, it is evident that there is an inwards force (the normal reaction force) that is holding them in a circular path. If the walls disintegrated and this normal force ceased to act, they would not 'fly outwards' but move at a tangent to their circle.

The Gravitron can rotate at 24 rpm and has a radius of 7 m. The centripetal acceleration can be over 40 m s^{-2}. This is caused by a very large centripetal force from the wall i.e. the normal force, F_N, which is greater than the weight force, F_g. Since the wall exerts such a large force, the patrons are pinned firmly to the wall as an upwards frictional force, F_f, acts to hold them up. The floor then drops away. It is important to note that there is no outwards force acting. In fact, as you can see in Figure 5.2.8, these forces are unbalanced and the net force is equal in size and direction to the normal force towards the centre of the circle.

FIGURE 5.2.7 There is a large inwards force from the wall (a normal reaction force) that causes these children to travel in a circular path.

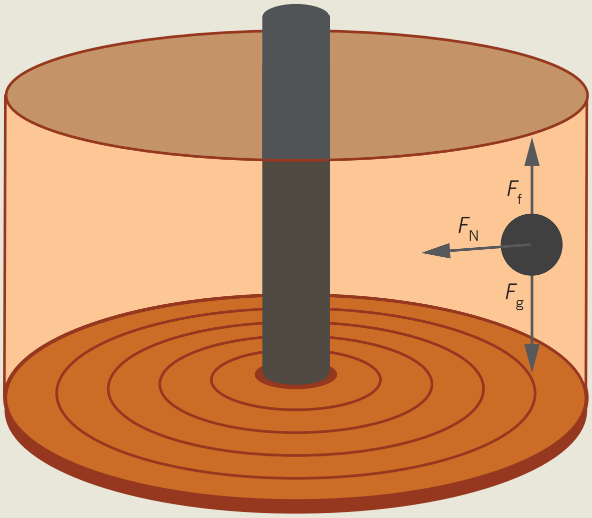


FIGURE 5.2.8 The forces acting on the person are unbalanced. There is an unbalanced force from the wall, F_N, giving the person a centripetal acceleration.

BALL ON A STRING

You may have played Totem Tennis at one time. This is a game where a ball is attached to a pole by a string and can travel in a horizontal circle, although the string itself is not horizontal. This kind of motion is shown in Figure 5.2.9.

If the ball at the end of the string was swinging slowly, the string would swing down at an angle closer to the pole. If the ball was swung faster, the string would become closer to being horizontal. In fact, it is not possible for the string to be absolutely horizontal, although as the speed increases, the closer to horizontal it becomes. This system is known as a conical pendulum.

If the angle of the conical pendulum is known, trigonometry can be used to find the radius of the circle and the forces involved.

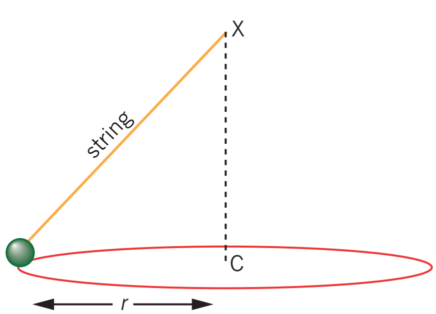
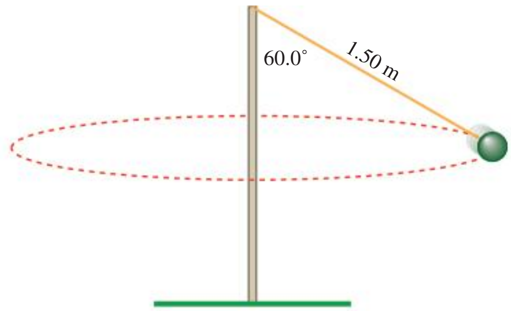


FIGURE 5.2.9 This ball is travelling in a horizontal circular path of radius r. The centre of its circular motion is at C.

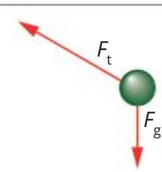
Worked example 5.2.3

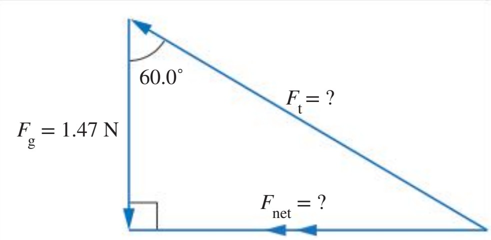
OBJECT ON THE END OF A STRING

During a game of Totem Tennis, the ball of mass 150 g is swinging freely in a horizontal circular path. The cord is 1.50 m long and is at an angle of 60.0° to the vertical, as shown in the diagram.



a Calculate the radius of the ball's circular path.	
Thinking	**Working**
The centre of the circular path is not the top end of the cord, but is where the pole is level with the ball. Use trigonometry to find the radius.	r = 1.50 sin 60.0° = 1.30 m

b Draw and identify the forces that are acting on the ball at the instant shown in the diagram.	
Thinking	**Working**
There are two forces acting—the tension in the cord, F_t, and gravity, F_g. These forces are unbalanced.	

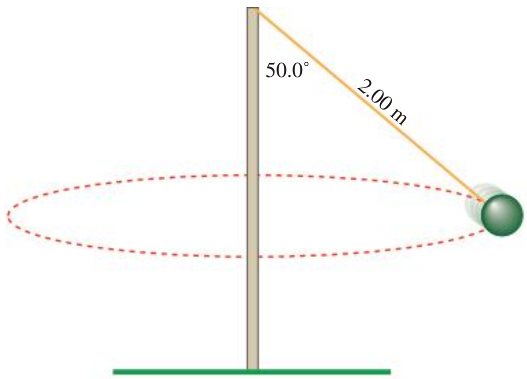
c Determine the net force that is acting on the ball at this time.	
Thinking	**Working**
First calculate the weight force, F_g.	$F_g = mg$ = 0.150 × 9.8 = 1.47 N
The ball has an acceleration that is towards the centre of its circular path. This is horizontal and towards the left at this instant. The net force will also lie in this direction at this instant. A force triangle and trigonometry can be used here.	 F_{net} = 1.47 tan 60.0° = 2.55 N towards the left

d Calculate the size of the tensile force in the cord.	
Thinking	**Working**
Use trigonometry to find F_t.	$F_t = \frac{1.47}{\cos 60.0°}$ = 2.94 N

Worked example: Try yourself 5.2.3

OBJECT ON THE END OF A STRING

During a game of Totem Tennis, the ball of mass 200 g is swinging freely in a horizontal circular path. The cord is 2.00 m long and is at an angle of 50.0° to the vertical, as shown in the diagram.



a Calculate the radius of the ball's circular path.

b Draw and identify the forces that are acting on the ball at the instant shown in the diagram.

c Determine the net force that is acting on the ball at this time.

d Calculate the size of the tensile force in the cord.

5.2 Review

SUMMARY

- Frequency, f, is the number of revolutions each second and is measured in hertz (Hz).
- Period, T, is the time for one revolution and is measured in seconds.
- The relationship between T and f is:
  $f = \frac{1}{T}$ and $T = \frac{1}{f}$
- An object moving with a uniform speed in a circular path of radius, r, and with a period, T, has an average speed that is given by:
  $v = \frac{2\pi r}{T}$
- The velocity of an object moving (with a constant speed) in a circular path is continually changing. The velocity vector is always directed at a tangent to the circular path.
- An object moving in a circular path (with a constant speed) has an acceleration due to its circular motion. This acceleration is directed towards the centre of the circular path and is called centripetal acceleration, a:
  $a = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2}$
- Centripetal acceleration is a consequence of a centripetal force acting to make an object move in a circular path.
- Centripetal forces are directed towards the centre of the circle and their magnitude can be calculated by using Newton's second law:
  $F_{net} = \frac{mv^2}{r} = \frac{4\pi^2 rm}{T^2}$
- Centripetal force is always supplied by a real force, the nature of which depends on the situation. The real force is commonly friction, gravitation or the tension in a string or cable.

5.2 Review *continued*

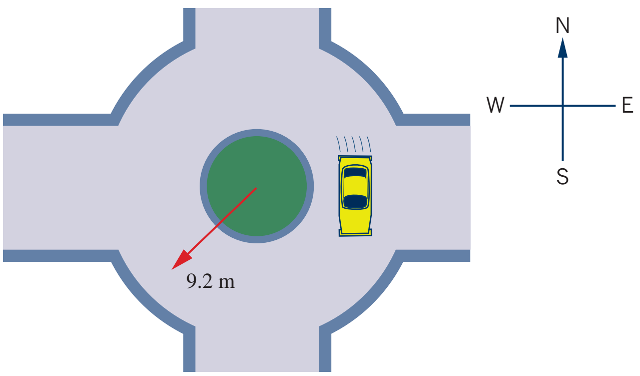
KEY QUESTIONS

1 A car is travelling with a constant speed around a roundabout. What is the centripetal force that is causing this circular motion?
- **A** gravity
- **B** friction
- **C** drag
- **D** tension

2 A boy is swinging a yo-yo in a horizontal circle 5 times each second. What is the period of the yo-yo?

The following information applies to questions 3–7.

A car of mass 1200 kg is travelling on a roundabout in a circular path of radius 9.2 m. The car moves with a constant speed of 8.0 m s^{-1}. The direction of the car is anticlockwise around the roundabout when viewed from above as shown.



3 Which two of the following statements correctly describe the motion of the car as it travels around the roundabout? More than one answer is possible.
- **A** It has a constant speed.
- **B** It has a constant velocity.
- **C** It has zero acceleration.
- **D** It has an acceleration that is directed towards the centre of the roundabout.

4 When the car is in the position shown in the diagram what is the:
- **a** speed of the car
- **b** velocity of the car
- **c** magnitude and direction of the acceleration of the car?

5 Calculate the magnitude and direction of the net force acting on the car at the position shown.

6 Sometime later, the car has travelled halfway around the roundabout. What is the:
- **a** velocity of the car at this point
- **b** direction of its acceleration at this point?

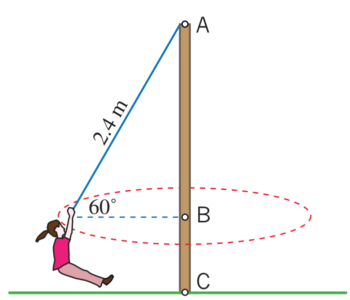
7 If the driver of the car kept speeding up, what would eventually happen to the car as it travelled around the roundabout? Explain your answer.

8 An ice skater of mass 50 kg is skating in a horizontal circle of radius 1.5 m at a constant speed of 2.0 m s^{-1}. Answer the questions below about the ice skater's motion.
- **a** Determine the magnitude of the skater's acceleration.
- **b** Are the forces acting on the skater balanced or unbalanced? Explain.
- **c** Calculate the magnitude of the centripetal force acting on the skater.

9 Fiona and Mark are flying their remote-controlled model plane. It has a mass of 1.6 kg and travels in a horizontal circular path of radius 62 m with a speed of 50 km h^{-1}. A radio transmitter controls the plane so there are no strings attached. Answer the questions below about the plane's motion.
- **a** Calculate the period of the model plane's motion.
- **b** Determine the magnitude of the net force that is acting on the plane.

10 An athlete competing at a junior sports meet swings a 2.5 kg hammer in a horizontal circle of radius 0.80 m at 2.0 revolutions per second. Assume that the wire is horizontal at all times.
- **a** What is the period of rotation of the ball?
- **b** What is the orbital speed of the ball?
- **c** What is the magnitude of the acceleration of the ball?
- **d** What is the magnitude of the net force acting on the ball?

11 A child of mass 30 kg is playing on a maypole swing in a playground. The rope is 2.4 m long and at an angle of 60° to the horizontal as she swings freely in a circular path. Ignore the mass of the rope in your calculations.



- **a** Calculate the radius of her circular path.
- **b** Identify the forces that are acting on her as she swings freely.
- **c** What is the direction of her acceleration when she is at the position shown in the diagram?
- **d** Calculate the net force acting on the girl.
- **e** What is her speed as she swings?

5.3 Circular motion on banked tracks

The previous section focused on relatively simple situations involving uniform circular motion in a horizontal plane. However, there are more complex situations involving this type of motion. On many road bends, the road is not horizontal, but is at a small angle to the horizontal. This enables vehicles to travel at higher speeds without skidding. A more dramatic example of this effect is at a cycling velodrome like that shown in Figure 5.3.1. The Darebin International Sports Centre in Thornbury has a velodrome that has banked or inclined corners that peak at 43°. This enables the cyclists to travel at much higher speeds than if the track were flat. This section examines the principles underlying banked cornering and applies Newton's laws to solving problems involving circular motion on banked tracks.

FIGURE 5.3.1 The Australian women's pursuit track cycling team in action on a banked velodrome track during the London Olympics in 2012.

BANKED CORNERS

Cars and bikes can travel much faster around corners when the road or track surface is inclined or banked at some angle to the horizontal. **Banked tracks** are most obviously used at cycling velodromes or motor sport events such as NASCAR races. Road engineers also design roads to be banked in places where there are sharp corners such as exit ramps on freeways.

When cars travel in circular paths on horizontal roads, they are relying on the force of friction between the tyres and the road to provide the sideways force that keeps the car turning in the circular path.

Consider a car travelling clockwise around a horizontal roundabout with a constant speed, v. As can be seen in Figure 5.3.2, the car has an acceleration towards C (the centre of the circle) and so the net force is also sideways on the car towards C.

The forces acting on the car are shown in Figure 5.3.3. As you can see, the vertical forces (gravity and the normal reaction force) are balanced. The only horizontal force is the sideways force that the road exerts on the car tyres. This is a force of friction, F_f, and is unbalanced, so this is equal to the net force, F_{net}.

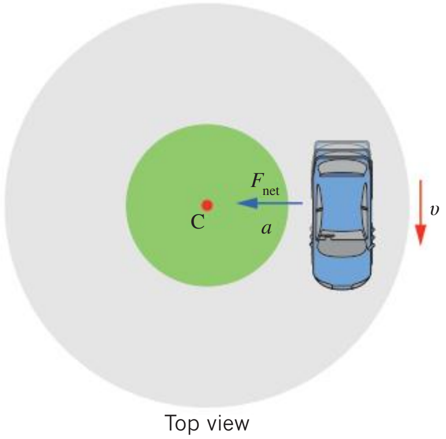


FIGURE 5.3.2 The car is travelling in a circular path on a horizontal track.

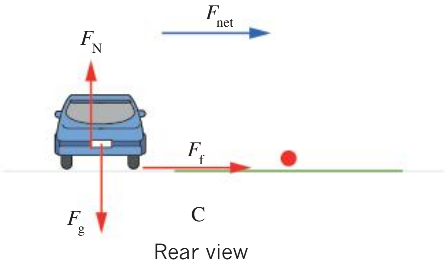


FIGURE 5.3.3 The vertical forces balance, and it is friction between the tyres and the road that enables the car to corner.

If the car drove over an icy patch, there would be no friction and the car would not be able to turn. It would skid in a straight line at a tangent to the circular path.

Banking the road reduces the need for a sideways frictional force and allows cars to travel faster without skidding off the road and away from the circular path. Consider the same car travelling around a circular, banked road with constant speed, v, as shown in Figure 5.3.4. It is possible for the car to travel at a speed so that there is no sideways frictional force. This is called the **design speed** and it is dependent on the angle, θ, at which the track is banked. At this speed, the car exhibits no tendency to drift higher or lower on the track.

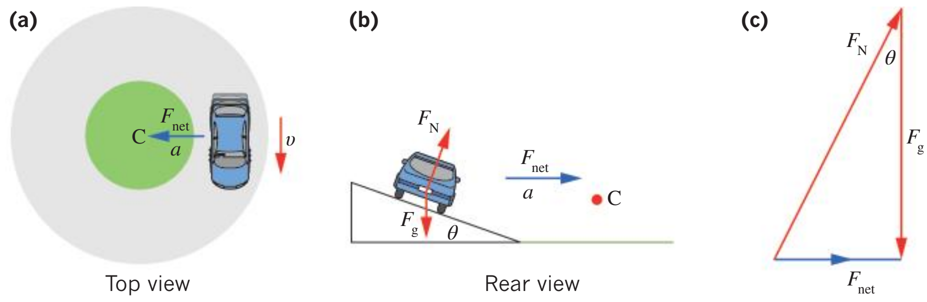


FIGURE 5.3.4 (a) The car is travelling in a circular path on a banked track. (b) The acceleration and net force are towards C. The banked track means that the normal force (F_N) has an inwards component. This is what enables the car to turn the corner. (c) Vector addition gives the net force (F_{net}) as acting horizontally towards the centre.

The car still has an acceleration towards the centre of the circle, C, and so there must be an unbalanced force in this direction. Due to the banking, there are now only two forces acting on the car: its weight, F_g, and the normal force, F_N, from the track.

As can be seen in part (b) of Figure 5.3.4, these forces are unbalanced. They add together to give a net force that is horizontal and directed towards C.

At the design speed, the angle of bank, θ, of the road or track can be found by using:

$$\tan\theta = \frac{F_{net}}{F_g}$$

where F_{net} is the force acting to the centre of the circle (N)

F_g is the force due to gravity on the object (N).

Extending this equation by substituting $F_{net} = \frac{mv^2}{r}$ and $F_g = mg$ gives:

$$\tan\theta = \frac{v^2}{rg} \text{ and hence } \theta = \tan^{-1}\left(\frac{v^2}{rg}\right)$$

where m is the mass of the vehicle and passengers (kg)

v is the speed of the vehicle (m s^{-1})

r is the radius of the track (m)

θ is the angle of bank (degrees)

g is the acceleration due to gravity (9.8 m s^{-2} near the surface of the Earth)

If the angle and weight are known, trigonometry can be used to calculate the net force (see Figure 5.3.4(c)) and therefore the design speed.

Rearranging $\tan\theta = \frac{v^2}{rg}$ to make the design speed, v, the subject gives:

$$v^2 = rg\tan\theta$$

$$v = \sqrt{rg\tan\theta}$$

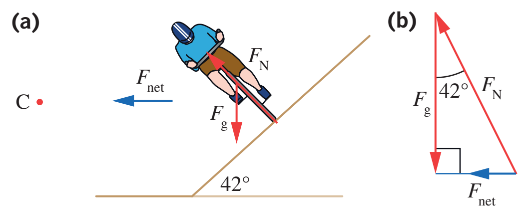
It is worth noting that the normal force will be larger here than on a flat track. In the case of a cyclist, the rider and bike would feel a larger force acting from the road when they are on a banked track compared to when they are cycling on a flat track (see Figure 5.3.5).

FIGURE 5.3.5 Australian cyclist Anna Meares on this banked velodrome track is cornering at speeds far higher than she could use on a flat track. Cyclists on a velodrome do not need to rely on friction to turn, and so experience a larger normal force than usual.

Worked example 5.3.1

BANKED CORNERS

A curved section of track on an Olympic velodrome has a radius of 50 m and is banked at an angle of 42° to the horizontal. A cyclist of mass 75 kg is riding on this section of track at the design speed.

<table>
<tr><td colspan="2">a Calculate the net force acting on a cyclist at this instant if they are riding at the design speed.</td></tr>
<tr><td>Thinking</td><td>Working</td></tr>
<tr><td>Draw a force diagram and include all forces acting on the cyclist.</td><td>The forces acting on the cyclist are gravity and the normal force from the track, and these are unbalanced. The net force is horizontal and towards the centre of the circular track as shown in diagram (a) and the force triangle of diagram (b).

</td></tr>
<tr><td>Calculate the weight force, F_g.</td><td>$F_g = mg$
$= 75 \times 9.8$
$= 735$ N</td></tr>
<tr><td>Use the force triangle and trigonometry to work out the net force, F_{net}.</td><td>$\tan\theta = \frac{F_{net}}{F_g}$
$\tan 42° = \frac{F_{net}}{735}$
$F_{net} = 0.90 \times 735$
$= 662$ N</td></tr>
<tr><td>As force is a vector, a direction is needed in the answer.</td><td>Net force is 662 N horizontally towards the centre of the circle.</td></tr>
</table>

<table>
<tr><td colspan="2">b Calculate the design speed for this section of the track.</td></tr>
<tr><td>Thinking</td><td>Working</td></tr>
<tr><td>List the known values.</td><td>$m = 75$ kg
$r = 50$ m
$\theta = 42°$
$F_g = 735$ N
$F_{net} = 662$ N
$v = ?$</td></tr>
<tr><td>Use the design speed formula.</td><td>$v = \sqrt{rg\tan\theta}$
$= \sqrt{50 \times 9.8 \times \tan 42°}$
$= 21$ m s^{-1}</td></tr>
</table>

PHYSICSFILE

Wall of Death

In some amusement parks in other parts of the world, there is a ride known menacingly as the Wall of Death (see Figure 5.3.6). It consists of a cylindrical enclosure with vertical walls. People on bicycles and motorbikes ride into the enclosure and around the vertical walls, so the angle of banking is 90°! The riders need to keep moving and are depending on friction to hold them up. By travelling fast, the centripetal force (the normal force from the wall) is large and this increases the size of the grip (friction) between the wall and tyres. If the rider slammed on the brakes and stopped, they would simply plummet to the ground.

FIGURE 5.3.6 For a rider to successfully conquer the Wall of Death, they need to travel reasonably fast and there must be good grip between the tyres and the track. The rider is relying on friction to maintain their motion along the wall.

Worked example: Try yourself 5.3.1

BANKED CORNERS

A curved section of track on an Olympic velodrome has radius of 40 m and is banked at an angle of 37° to the horizontal. A cyclist of mass 80 kg is riding on this section of track at the design speed.

a Calculate the net force acting on a cyclist at this instant as they are riding at the design speed.

b Calculate the design speed for this section of the track.

EXTENSION

Leaning into corners

In many sporting events, participants need to travel around corners at high speeds. As shown in Figure 5.3.7, motorbike riders lean their bikes over almost onto the track as they corner. This leaning technique is also evident in ice skating, bicycle races, skiing and even when you run round a corner. It enables the competitor to corner at high speed without falling over.

FIGURE 5.3.7 Australia's Casey Stoner won the 2012 Moto GP championship. Here he is leaning his bike as he takes a corner at Phillip Island. Leaning into the corner enables him to corner at higher speeds. In fact, the bike would go out of control if he did not lean it.

Consider a bike rider cornering on a horizontal road surface (see Figure 5.3.8). The forces acting on the bike and rider are unbalanced. The forces are the weight force, F_g, and the force from the track. The track exerts a reaction force, on the rider that acts both inwards and upwards. The inwards component is the frictional force, F_f, between the track and the tyres. The upwards component is the normal force, F_N, from the track.

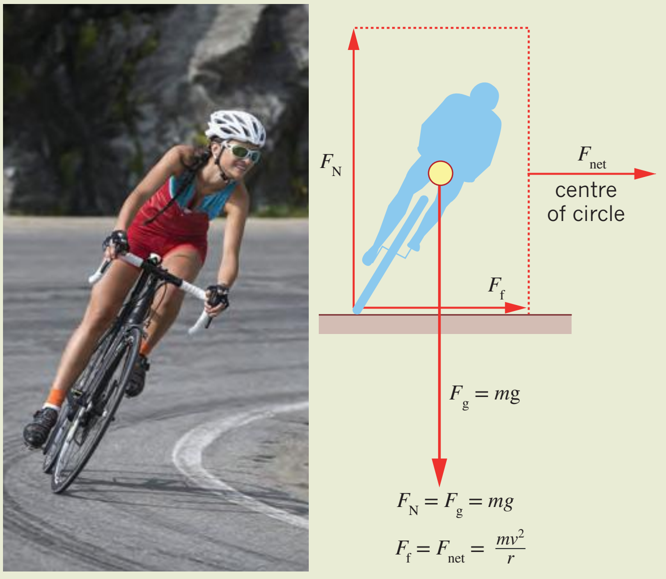


FIGURE 5.3.8 The forces acting as the rider turns a corner are the weight, F_g, the normal force, F_N, and the friction, F_f, between the tyres and the road. The friction supplies the unbalanced force that leads to the corner turning motion.

The rider is travelling in a horizontal circular path at constant speed, and so has a centripetal acceleration directed towards the centre of the circle. Therefore, the net force is directed towards the centre of the circle. By analysing the vertical and horizontal components in Figure 5.3.8, you see that the weight force, F_g, must balance the normal force, F_N. The net force that is producing the centripetal acceleration is supplied by the frictional force, F_f. In other words, the rider is depending on a sideways frictional force to turn the corner. An icy or oily patch on the track would cause the tyres to slide out from under the rider, and he or she would slide painfully along the road at a tangent to the circular path.

5.3 Review

SUMMARY

- A banked track is one where the track is inclined at some angle to the horizontal. This enables vehicles to travel at higher speeds when cornering, compared with around a horizontal curved path.
- Banking a track eliminates the need for a sideways frictional force to turn. When the speed and angle are such that there is no sideways frictional force, the speed is known as the design speed.
- The forces acting on a vehicle travelling at the design speed on a banked track are gravity and the normal force from the track. These forces are unbalanced and add to give a net force directed towards the centre of the circular motion.
- At the design speed, the angle of bank of the track, $\theta = \tan^{-1}\left(\frac{v^2}{rg}\right)$.
- For a given bank angle and curve radius, the design speed is given by: $v = \sqrt{rg \tan \theta}$.

KEY QUESTIONS

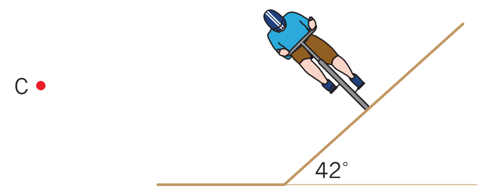
1 A cyclist is riding along a circular section of a velodrome where the radius is 30 m and the track is inclined at 30° to the horizontal. The cyclist is riding at the design speed and maintains a constant speed. Describe the direction of the acceleration on the cyclist.

2 An architect is designing a velodrome and the original plans have semi-circular sections of radius 15 m and a banking angle of 30°. The architect is asked to make changes to the plans that will increase the design speed for the velodrome. What two design elements could the architect change in order to meet this requirement?

3 A racing car is travelling around a circular banked track which has a design speed of 100 km h^{-1}. On one lap, the car travels at 150 km h^{-1}. At this higher speed, the car would tend to travel in a different position along the banked surface. Would the car travel higher or lower up the banked track? Explain your answer.

4 A racing car travels at high speed along a horizontal track and tries to turn a corner. The car skids and loses control. The racing car then travels along a banked track and is able to travel much faster around the corners without skidding at all. Complete the sentences below by choosing the correct term in bold.
On the horizontal track, the car is depending on the force of **friction**/**weight** to turn the corner. The **friction**/**normal** force is equal to the **weight**/**friction** of the car so these vertical forces are **balanced**/**unbalanced**. When driving on the banked track, the **normal**/**weight** force is not vertical and so is not balanced by the **weight**/**normal** force. In both cases, the forces acting on the car are unbalanced.

5 Copy and complete the following diagram by drawing and labelling the normal force, weight force and net force acting on the bicycle.



The following information relates to questions 6 and 7.
A cycling velodrome has a turn that is banked at 33° to the horizontal. The radius of the track at this point is 28 m.

6 Determine the speed (in km h^{-1}) at which a cyclist of mass 55 kg would experience no sideways force on their bike as they ride this section of track.

7 **a** Calculate the size of the normal force that is acting on the cyclist.
b How would this compare with the normal force if they were riding on a flat track?

8 A car racing track is banked so that when the cars corner at 40 m s^{-1}, they experience no sideways frictional forces. The track is circular with a radius of 150 m.
Calculate the angle to the horizontal at which the track is banked.

The following information applies to questions 9 and 10.
A section of track at a NASCAR raceway is banked to the horizontal. The track section is circular with a radius of 80 m and it has a design speed of 18 m s^{-1}. A car of mass 1200 kg is being driven around the track at 18 m s^{-1}.

9 **a** Calculate the magnitude of the net force acting on the car (in kN).
b Calculate the angle to the horizontal at which the track is banked.

10 The driver now drives around the track at 30 m s^{-1}. What would the driver have to do to maintain their circular path around the track?

5.4 Circular motion in a vertical plane

In previous sections, the motion of objects travelling in circular paths was discussed. It was explained that a body moving with constant speed in a horizontal circular path has an acceleration that is directed towards the centre of the circle. The same applies for vertical circular paths.

If you have ever been on a rollercoaster ride, you will have travelled over humps and down through dips at high speeds and, at times, in circular arcs (see Figure 5.4.1). There are also other rides that travel through full 360° vertical circles. During these rides, your body will experience forces that you may or may not find pleasant.

When you travel on a rollercoaster, you can experience quite strong forces pushing you down into the seat as you fly through the dips. On the other hand, as you travel over the humps, you tend to lift off the seat. These forces relating to circular motion in a vertical plane will be discussed in this section. Like in the previous sections, Newton's laws are used to solve problems involving this type of circular motion.

FIGURE 5.4.1 This rollercoaster has a circular path in a vertical plane at this point.

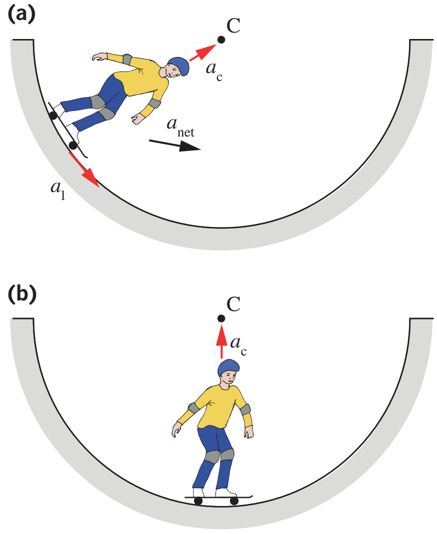


FIGURE 5.4.2 (a) High on the sides of the 'pipe', the skateboarder speeds up, and so has both a linear and a centripetal acceleration. The net acceleration, a_{net}, is not towards C. (b) At the lowest point the speed of the skateboarder is momentarily constant, so there is no linear acceleration. The acceleration is supplied completely by the centripetal acceleration, a_c, and is acting towards C.

MOVING IN VERTICAL CIRCLES

A body moving with constant speed in a horizontal circular path has an acceleration that is directed towards the centre of the circle. The same applies for vertical circular paths. However, circular motion in a vertical plane in real life is often more complex as it does not usually involve constant speeds.

An example of this is illustrated in Figure 5.4.2(a). The speed of the skateboarder practising in a half-pipe will increase on the way down as gravitational potential energy is converted into kinetic energy. This means the skater will experience a linear acceleration, a_l, as well as a centripetal acceleration, a_c. The resultant acceleration is not directed towards the centre of the circular path.

At the bottom of the 'pipe', the skateboarder will be neither slowing down nor speeding up, so the acceleration is purely centripetal at this point, as shown in part (b) of Figure 5.4.2. The same applies at the very top of a circular path. For this reason, motion at these points is easier to analyse.

Uniform horizontal motion

Theme park rides make you appreciate that the forces you experience throughout a ride can vary greatly. First, consider the case of a person in a rollercoaster cart, like that shown in Figure 5.4.3, travelling horizontally at 4.0 m s^{-1}. If the person's mass is 50 kg and the gravitational field strength is 9.8 m s^{-2}, the forces acting on the person can be calculated. These forces are the weight, F_g, and the normal reaction force, F_N, from the seat.

The person is moving in a straight line with a constant speed, so there is no unbalanced force acting. The weight force balances the normal reaction force from the seat. The normal force is therefore 490 N up, which is what usually acts upwards on this person when moving horizontally and they would feel the same as their usual weight.

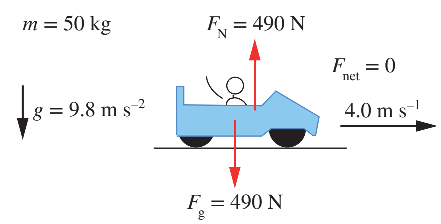


FIGURE 5.4.3 The vertical forces are in balance in this situation, i.e. $F_N = F_g$.

Circular motion: travelling through dips

Now consider the forces that act on the person after the cart has reached the bottom of a circular dip of radius 2.5 m and is moving at 8.0 m s^{-1}. Figure 5.4.4 illustrates these forces.

The person will have a centripetal acceleration due to the circular path. This centripetal acceleration is directed towards the centre, C, of the circular path—in this case, vertically upwards. The person's centripetal acceleration, a, is:

$$a = \frac{v^2}{r}$$
$$= \frac{8.0^2}{2.5}$$
$$= 26 \text{ m s}^{-2} \text{ towards C, or upwards}$$

The net (centripetal) force acting on the person is given by:

$$F_{net} = ma$$
$$= 50 \times 26$$
$$= 1300 \text{ N upwards}$$

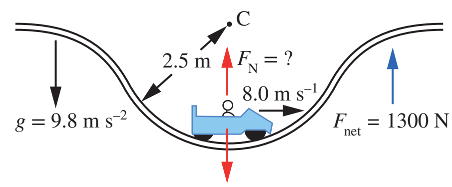


FIGURE 5.4.4 The person has a centripetal acceleration that is directed upwards towards the centre of the circle, and so the net force is also upwards. In this case, the magnitude of the normal force, F_N, is greater than the weight, F_g, and produces a situation where the rider feels heavier than usual.

The normal force, F_N, and the weight force, F_g, are not in balance anymore. They add together to give an upwards force of 1300 N. This indicates that the normal force must be greater than the weight force by 1300 N. In other words, the normal force is 490 N + 1300 N = 1790 N up. This is over three times larger than the normal force of 490 N that usually acts. That is the reason why, when in a ride, you feel the seat pushing up against you much more strongly at this point. This normal force of 1790 N in this instance is equal to the apparent weight of the person and indicates they would feel much heavier than usual.

Circular motion: travelling over humps

Now consider the situation as the cart moves over the top of a hump of radius 2.5 m with a lower speed of 2.0 m s^{-1}, as illustrated in Figure 5.4.5.

The person now has a centripetal acceleration that is directed vertically downwards towards the centre, C, of the circle. Therefore, the net force acting at this point is directed vertically downwards. The centripetal acceleration is:

$$a = \frac{v^2}{r}$$
$$= \frac{2.0^2}{2.5}$$
$$= 1.6 \text{ m s}^{-2} \text{ towards C, or downwards}$$

The net (centripetal) force is:

$$F_{net} = ma$$
$$= 50 \times 1.6$$
$$= 80 \text{ N downwards}$$

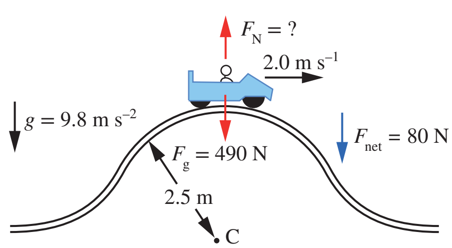


FIGURE 5.4.5 The centripetal acceleration is downwards towards the centre of the circle, and so the net force is also in that direction. At this point, the magnitude of the normal force, F_N, is less than the weight, F_g, of the person.

As in the dip, the weight force and the normal force are not in balance. They add to give a net force of 80 N down. The weight force, F_g, must therefore be 80 N greater than the normal force, F_N. This tells us that the normal force is:

490 N + (–80) N = 410 N up.

How the normal force varies during the ride

It is interesting to compare the normal forces that act on the person in these three situations.

- The normal force when travelling horizontally is 490 N upwards.
- At the bottom of the dip, the normal force is 1790 N upwards. In other words, in the dip, the seat pushes into the person with a greater force than usual. This gives the person an apparent weight of 1790 N and makes the person feel much heavier than normal. If the person had been sitting on weighing scales at this time, it would have shown a higher than usual reading.
- At the top of the hump, the normal force is 410 N upwards. In other words, over the hump, the seat pushes into the person with a smaller force than usual. This gives the person an apparent weight of 410 N and gives them the sensation of feeling lighter.

The weight of the person has not changed. F_g is 490 N throughout the duration of the ride; it is the normal force acting on them that varies. The normal force is equal to the person's apparent weight, and this makes the person 'feel' heavier and lighter as they travel through the dips and humps respectively.

PHYSICSFILE

Fighter pilots

A fighter pilot in a vertical loop manoeuvre can safely experience centripetal accelerations of up to around 5g, or 49 m s^{-2}. In a loop where the g-forces are greater than this, the pilot may pass out. If the pilot flies with his or her head inside the loop, the centripetal acceleration of the plane will push the pilot into their seat and make the blood flow away from their head. The resulting lack of blood in the brain may cause the pilot to 'grey out' and they may totally lose consciousness ('black out'). This type of force is a called a positive-g force. Fighter pilots wear 'g-suits', which pressurise the legs to prevent blood flowing into them which helps them to maintain consciousness (see Figure 5.4.6).

On the other hand, if the pilot's head is on the outside of the loop, the centripetal acceleration will pull the pilot onto their harness and the additional blood flow to the head can make the whites of the eyes turn red. The excess blood flow in the head may cause 'red out'. This type of force is called a negative-g force.

FIGURE 5.4.6 Fighter pilots wear pressurised suits to allow their bodies to withstand the large forces that act during tight turns.

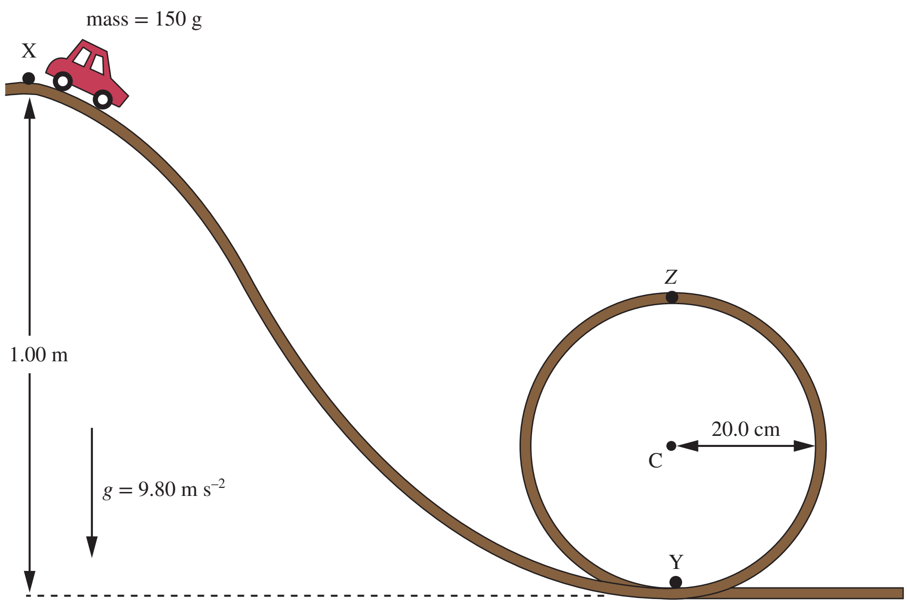
A rollercoaster cart going through dips and over humps is always moving above the rollercoaster track, so the normal reaction force acts upwards. It is also possible for the cart to travel on the underside of the track when it goes upside down through a loop. Worked example 5.4.1 shows a toy car travelling through a vertical loop, on the inside of the loop.

Worked example 5.4.1

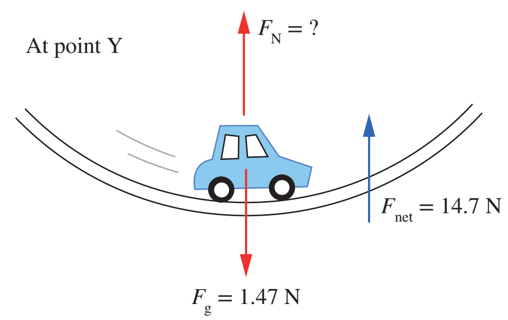
VERTICAL CIRCULAR MOTION

A student arranges a toy car track with a vertical loop of radius 20.0 cm, as shown.

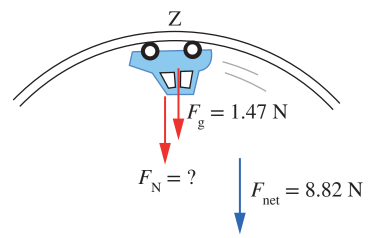
A toy car of mass 150 g is released from rest at a height of 1.00 m at point X. The car rolls down the track and travels inside the loop. Assume g is 9.80 m s^{-2}, and ignore friction.



a Calculate the speed of the car as it reaches the bottom of the loop, point Y.	
Thinking	**Working**
Note all the variables given to you in the question.	At X: m = 150 g = 0.150 kg h = 1.00 m v = 0 g = 9.80 m s^{-2}
Use an energy approach to calculate the speed. Calculate the total mechanical energy first.	The initial speed is zero, so E_k at X is zero. The total mechanical energy, E_m, at X is: $E_m = E_k + E_g$ $= \frac{1}{2}mv^2 + mg\Delta h$ $= 0 + (0.150 \times 9.80 \times 1.00)$ = 1.47 J
Use conservation of energy ($E_m = E_k + E_g$) to determine the velocity at point Y. As the car rolls down the track, it loses its gravitational potential energy and gains kinetic energy. At the bottom of the loop (Y), the car has zero potential energy.	At Y: E_m = 1.47 J h = 0 E_g = 0 $E_m = E_k + E_g$ $E_m = \frac{1}{2}mv^2 + mg\Delta h$ $1.47 = 0.5 \times 0.150v^2 + 0$ $v^2 = \frac{1.47}{0.0750}$ $v = \sqrt{19.6}$ = 4.43 m s^{-1}

b Calculate the normal reaction force from the track at point Y.	
Thinking	**Working**
To solve for F_N, start by working out the net, or centripetal, force. At Y, the car has a centripetal acceleration towards C (i.e., upwards), so the net (centripetal) force must also be vertically upwards at this point.	$F_{net} = \frac{mv^2}{r}$ $= \frac{0.150 \times 4.43^2}{0.200}$ $= 14.7$ N up
Calculate the weight force, F_g, and add it to a force diagram.	$F_g = mg$ $= 0.150 \times 9.80$ $= 14.7$ N down 
Work out the normal force using vectors. Note up as positive and down as negative for your calculations. The forces acting are unbalanced, as the car has a centripetal acceleration upwards (towards C). The upwards (normal) force must be larger than the downwards force.	$F_{net} = F_g + F_N$ $+14.7 = -1.47 + F_N$ $F_N = +14.7 + 1.47$ $= 16.2$ N up Note that the force the track exerts on the car is much greater (by about ten times) than the weight force. If the car were travelling horizontally, the normal force would be just 1.47 N up.

c What is the speed of the car as it reaches point Z?	
Thinking	**Working**
Calculate the velocity from the values you have, using $E_m = E_k + E_g$.	At Z: $m = 0.150$ kg $\Delta h = 2 \times 0.200 = 0.400$ m Mechanical energy is conserved, so $E_m = 1.47$ J At Z: $E_m = E_k + E_g$ $= \frac{1}{2}mv^2 + mg\Delta h$ $1.47 = \frac{1}{2} \times 0.150 \times v^2 + 0.150 \times 9.80 \times 0.400$ $1.47 = 0.075v^2 + 0.588$ $0.075v^2 = 1.47 - 0.588$ $v^2 = 11.76$ $v = 3.43$ m s^{-1}

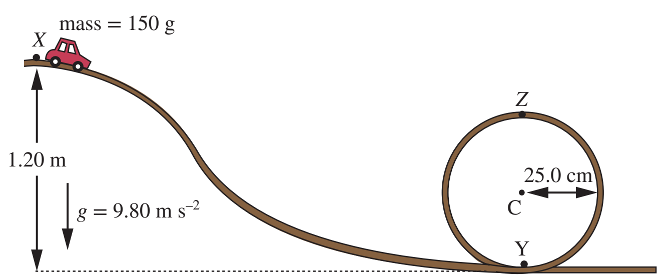
d What is the normal force acting on the car at point Z?	
Thinking	**Working**
To find F_N, start by working out the net, or centripetal, force. At Z, the car has a centripetal acceleration towards C (i.e., downwards), so the net (centripetal) force must also be vertically downwards at this point.	$F_{net} = \frac{mv^2}{r}$ $= \frac{0.150 \times 3.43^2}{0.200}$ = 8.82 N down
Work out the normal force using vectors. Note up as positive and down as negative for your calculations.	 $F_{net} = F_g + F_N$ $-8.82 = -1.47 + F_N$ $F_N = -8.82 + 1.47$ $= -7.35$ = 7.35 N down Note that there is still strong contact between the car and the track as given by the normal force, but that is only around half the size compared to the bottom of the track. If the car had progressively lower speeds, the normal force at Z would decrease and eventually drop to zero. At this point, the car would lose contact with the track, fall off the track and its acceleration would be equal to g.

Worked example: Try yourself 5.4.1

VERTICAL CIRCULAR MOTION

A student arranges a toy car track with a vertical loop of radius 25.0 cm, as shown.

A toy car of mass 150 g is released from rest at a height of 1.20 m at point X. The car rolls down the track and travels around the loop. Assume g is 9.80 m s^{-2}, and ignore friction for the following questions.



a Calculate the speed of the car as it reaches the bottom of the loop, point Y.

b Calculate the normal reaction force from the track at point Y.

c What is the speed of the car as it reaches point Z?

d What is the normal force acting on the car at point Z?

PHYSICS IN ACTION

Travelling upside down without falling out

You might have been on a rollercoaster like the one in Figure 5.4.7, where you were actually upside down at times during the ride. These rides use their speed and the radius of their circular path to prevent the riders from falling out. In theory, the safety harnesses worn by the riders are not needed to hold the people in their seats.

FIGURE 5.4.7 The thrill seekers on this rollercoaster ride don't fall out when upside down because the centripetal acceleration of the cart is greater than 9.8 m s^{-2} down.

The reason people don't fall out is that their centripetal acceleration while on the rollercoaster is greater than the acceleration due to gravity (9.8 m s^{-2}). To understand the significance of this, try the following activity. Stand up, reach up with one hand, place an eraser on the palm of that hand, then turn your hand palm down and move it rapidly towards the floor.

You should find, after one or two attempts, that it is possible to keep the eraser in contact with your hand as you 'push' it down. The eraser is upside down, but it is not falling out of your hand. Your hand must be for a short time moving down with a downwards acceleration in excess of 9.8 m s^{-2} and continually exerting a normal force on the eraser. This acceleration of 9.8 m s^{-2} down is the critical point in this exercise. If your hand had an acceleration less than this, the eraser would fall away from your hand to the floor. Try it to confirm that is what happens.

A similar principle holds with rollercoaster rides. The people on the ride don't fall out at the top because the motion of the rollercoaster gives them a centripetal acceleration that is greater than 9.8 m s^{-2} down. The engineers who designed the ride would have ensured that the rollercoaster moves with sufficient speed and in a circle of the appropriate radius so that this happens.

As an example, consider a ride of radius 15 m in a simple vertical circle (see Figure 5.4.8).

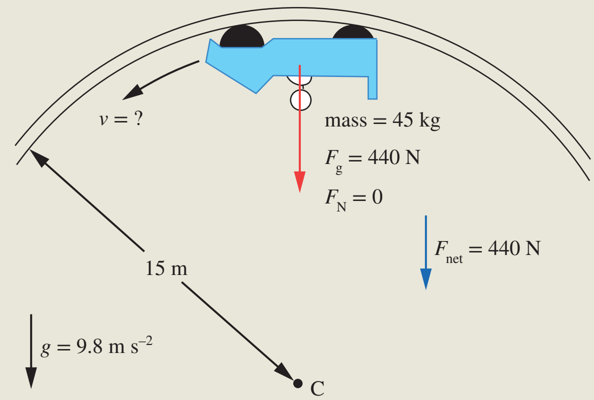


FIGURE 5.4.8 Rollercoaster travelling upside down through a loop. At the critical point where the rollercoaster just stays in contact with the track, the normal reaction force can be considered to be zero.

It is possible to calculate the speed that would ensure that a rider cannot fall out. Assume that the person has a mass of 45 kg and that g is 9.8 m s^{-2}. At the critical speed, the normal force, F_N, on the person will be zero. In other words, the seat will exert no force on the person at this speed. The centripetal force, F_{net}, is

$F_{net} = F_g + F_N$ but $F_N = 0$, so:

$= F_g$

therefore $\frac{mv^2}{r} = mg$

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

$= gr$

$v = \sqrt{gr}$

$= \sqrt{9.8 \times 15}$

$= 12 \text{ m s}^{-1}$

This speed is equal to 43 km h^{-1} and is the minimum needed to prevent the riders from falling out for that radius. In practice, the rollercoaster would move with a speed much greater than this to ensure that there was a significant force between the riders and their seats. Corkscrew rollercoasters can travel at up to 110 km h^{-1} and the riders can experience accelerations of up to 50 m s^{-2} or (5g). So, safety harnesses are really only needed when the speed is below the critical value; their primary function is to prevent people from moving around while on the ride.

5.4 Review

SUMMARY

- The gravitational force, F_g, and normal force, F_N, must be considered when analysing the motion of an object moving in a vertical circle.
- If the normal force is greater than the gravitational force ($F_N > F_g$) the passenger or rider will feel heavier.
- If the normal force is less than the gravitational force ($F_N < F_g$) the passenger or rider will feel lighter.
- In vertical circular motion, the gravitational force always acts vertically downwards regardless of position around the circle, the net force always acts towards the centre of the circle and the normal force always acts between the seat and the passenger or rider.
- The normal force and the gravitational force must add together as vectors in a force diagram to give the resultant as the net force.
- At the point where a moving object lifts off from its circular path, the normal force is zero. The object will be moving with a centripetal acceleration that is equal to that due to gravity (9.8 m s^{-2} down).
- Problems relating to motion in vertical circles can also be solved using an energy approach. This involves using the equation:

$$E_m = E_k + E_g = \frac{1}{2}mv^2 + mg\Delta h$$

KEY QUESTIONS

In the following questions, assume that g = 9.80 m s^{-2} and ignore the effects of air resistance.

The following information applies to questions 1 and 2.

A yo-yo is swung with a constant speed in a vertical circle.

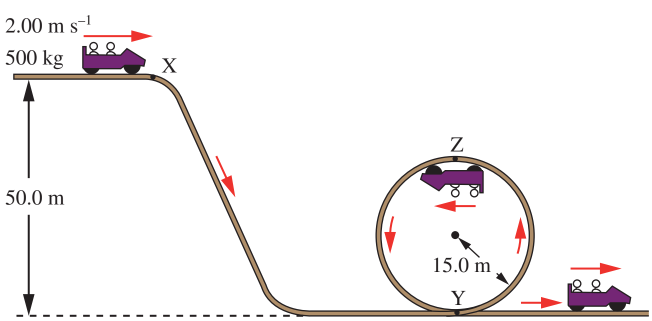
1 **a** Describe the magnitude of the acceleration of the yo-yo along its path.
b At which point in the circular path is there the greatest amount of tension in the string?
c At which point in the circular path is there the lowest amount of tension in the string?
d At which point is the string most likely to break?

2 If the yo-yo has a mass of 80 g and the radius of the circle is 1.5 m, find the minimum speed that this yo-yo must have at the top of the circle so that the cord does not slacken.

3 A car of mass 800 kg encounters a speed hump of radius 10 m. The car drives over the hump at a constant speed of 14.4 km h^{-1}.
a Name all the vertical forces acting on the car when it is at the top of the hump.
b Calculate the resultant force acting on this car when it is at the top of the hump.
c After travelling over the hump, the driver remarked to a passenger that she felt lighter as the car moved over the top of the speed hump. Is this possible? Explain your answer.
d What is the maximum speed (in km h^{-1}) that this car can have at the top of the hump and still have its wheels in contact with the road?

The following information applies to questions 4 and 5.

A popular amusement park ride is the 'loop-the-loop', in which a cart descends a steep incline at point X, enters a circular loop at point Y, and makes one complete revolution of the circular loop. The car, whose total mass is 500 kg, carries the passengers with a speed of 2.00 m s^{-1} when it begins its descent at point X from a vertical height of 50.0 m.



4 **a** Calculate the speed of the car at point Y.
b What is the speed of the car at point Z?
c Calculate the normal force acting on the car at Z.

5 What is the minimum speed that the car can have at point Z and still stay in contact with the track?

The following information applies to questions 6 and 7.

A stunt pilot appearing at an air show decides to perform a vertical loop so that she is upside down at the top of the loop. During the stunt she maintains a constant speed of 35 m s^{-1} while completing the 100 m radius loop.

6 Calculate the apparent weight of the 80 kg pilot when she is at the top of the loop.

5.4 Review *continued*

7 What minimum speed would the pilot need at the top of the vertical loop in order to experience zero normal force from the seat (i.e. to feel weightless)?

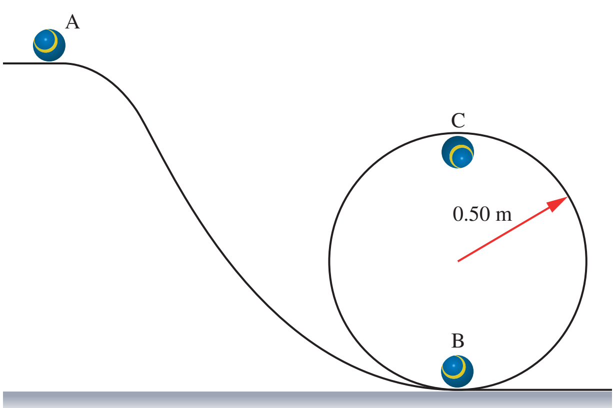
8 The maximum value of acceleration that the human body can safely tolerate for short time intervals is nine times that due to gravity. Calculate the maximum speed with which a pilot could safely pull out of a circular dive of radius 400 m.

9 A skateboarder of mass 55 kg is practising on a half-pipe of radius 2.0 m. At the lowest point of the half-pipe, the speed of the skater is 6.0 m s^{-1}. (Ignore air resistance and friction.)

a What is the acceleration of the skater at this point? Indicate both magnitude and direction.

b Calculate the size of the normal force acting on the skater at this point.

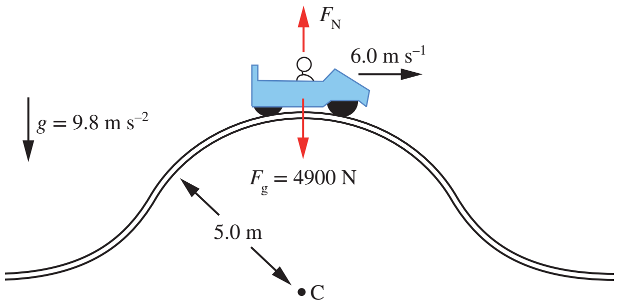
10 A ball bearing of mass 25 g is rolled along a smooth track in the shape of a loop-the-loop. The ball bearing is given a launch speed at A so that it just maintains contact with the track as it passes through point C. Ignore air resistance and friction.



a Determine the magnitude of the acceleration of the ball bearing as it passes point C.

b How fast is the ball bearing travelling at point C?

11 On the Mad Mouse ride, a cart of mass 500 kg encounters a hill of radius 5.0 m. The cart's speed at the top of the hump is 6.0 m s^{-1}.



a Calculate the magnitude and direction of the resultant force acting on this cart when it is at the top of the hump.

b Calculate the magnitude and direction of the normal force acting on this cart when it is at the top of the hump.

c What is the maximum speed that this cart can have at the top of the hump and still have its wheels in contact with the track?

5.5 Projectiles launched horizontally

A projectile is any object that is thrown or projected into the air and is moving freely—that is, it has no power source (such as a rocket engine or propeller) driving it. A netball as it is passed, a cricket ball that is hit for six, and an aerial skier flying through the air are all examples of projectiles. People have long argued about the path that projectiles follow; some thought they were circular or had straight sections. It is now known that if projectiles are not launched vertically and if air resistance is ignored, they move in smooth parabolic paths, like that shown in Figure 5.5.1. This section considers projectiles that are launched horizontally and uses Newton's laws to solve problems involving this type of motion.

FIGURE 5.5.1 A multi-flash photograph of a golf ball that has been bounced on a hard surface. The ball moves in a parabolic path.

PROJECTILE MOTION

It is a very common misconception that when a **projectile**, such as a netball, travels forwards through the air, it has a forwards force acting on it. This is incorrect. There may have been some forwards force acting as the projectile is launched, but once the projectile is released, this forwards force is no longer acting.

In fact, if air resistance is ignored, the only force acting on a projectile during its flight is its weight, which is the force due to gravity, F_g. This force is constant and always directed vertically downwards. This causes the projectile to continually deviate from a straight-line path to follow a parabolic path, as seen in Figure 5.5.2. This motion is known as **free fall**.

FIGURE 5.5.2 The motorcycle and rider are travelling in parabolic paths as they fly through the air.

Projectile motion is quite complex compared to straight-line motion. It must be analysed by considering the different components—horizontal and vertical—of the actual motion. The vertical and horizontal components of the motion are independent of each other and must be treated separately.

Given that the only force acting on a projectile is the force due to gravity, F_g, it follows that the projectile must have a vertical acceleration of 9.8 m s^{-2} downwards throughout its motion.

> In the vertical direction, a projectile accelerates due to the force of gravity, that is, at a rate of 9.8 m s^{-2} downwards.
>
> In the horizontal direction a projectile has a uniform velocity since there are no forces acting in this direction (if air resistance is ignored). So, the horizontal acceleration is zero.

PROJECTILES LAUNCHED HORIZONTALLY

Projectiles can be launched at any angle. The launch velocity needs to be resolved into vertical and horizontal components using trigonometry in order to complete most problems. For projectiles launched horizontally, calculating the vector components of the launch velocity is easy to do. That's because the initial vertical velocity is zero (but increases during the flight). The horizontal velocity is constant and is equal to the launch velocity. This can be verified using trigonometric ratios and a launch angle of 0°.

Tips for solving projectile motion problems

1. Construct a diagram showing the projectile's motion to set the problem out clearly. Write out the information supplied for the horizontal and vertical components separately.
2. In the horizontal direction the velocity, v, of the projectile is constant, so the only formula needed is $v_{av} = \frac{s}{t}$.
3. In the vertical direction, the projectile is moving with a constant acceleration (9.8 m s^{-2} down), and so the equations of motion for uniform acceleration must be used. These include:
   $v = u + at$
   $s = ut + \frac{1}{2}at^2$
   $v^2 = u^2 + 2as$
4. In the vertical direction it is important to clearly specify whether up or down is the positive or negative direction. Either choice will work just as effectively. The same convention needs to be used consistently throughout each problem.
5. If a projectile is launched horizontally, its horizontal velocity throughout the flight is the same as its initial velocity.
6. Pythagoras' theorem can be used to determine the actual speed of the projectile at any point.
7. If the velocity of the projectile is required, it is necessary to provide a direction with respect to the horizontal plane as well as the speed of the projectile.

PHYSICSFILE

Cartoon physics

It is easy to get the wrong idea about projectile motion when you watch cartoon characters running or driving off cliffs. In many cartoons, the character leaves the cliff and travels horizontally outward, stopping in mid-air (see Figure 5.5.3). Once they realise where they are, they immediately fall vertically downwards. This is not what happens in reality! The character should start falling in a smooth parabolic arc as soon as they leave the cliff-top!

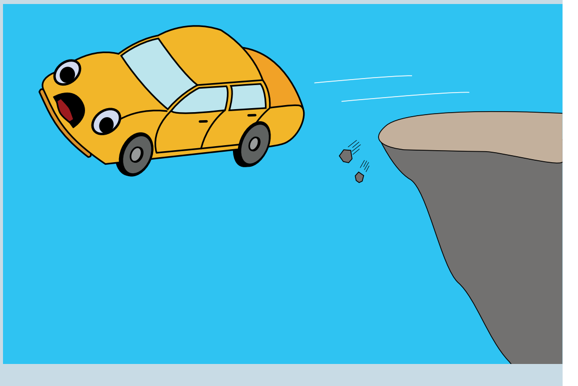
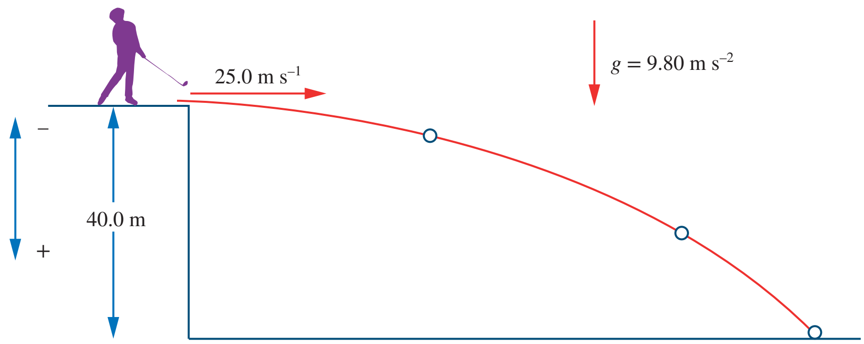


FIGURE 5.5.3 Many misconceptions can arise from what is shown in cartoons. In real life, this car would start falling as soon as it leaves the cliff top and it would travel in a parabolic arc.

Worked example 5.5.1

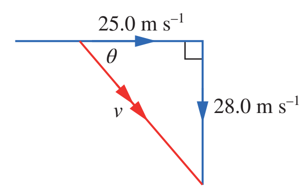
PROJECTILE LAUNCHED HORIZONTALLY

A golf ball of mass 150 g is hit horizontally from the top of a 40.0 m–high cliff with a speed of 25.0 m s^{-1}. Using g = 9.80 m s^{-2} and ignoring air resistance, calculate the following values.



a Calculate the time that the ball takes to land.	
Thinking	**Working**
Let the downward direction be positive. Write out the information relevant to the vertical component of the motion. Note that the instant the ball is hit, it is travelling only horizontally, so its initial vertical velocity is zero.	Down is positive. Vertically: u = 0 m s^{-1} s = 40.0 m a = 9.80 m s^{-2} t = ?
In the vertical direction, the ball has constant acceleration, so use equations for uniform acceleration. Select the equation that best fits the information you have.	$s = ut + \frac{1}{2}at^2$
Substitute values, rearrange and solve for t.	$40.0 = 0 + \frac{1}{2} \times 9.80 \times t^2$ $t = \sqrt{\frac{40.0}{4.90}}$ $= 2.86$ s

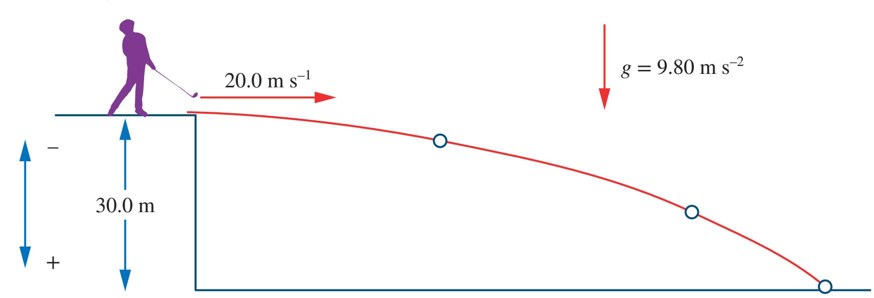
b Calculate the distance that the ball travels from the base of the cliff, i.e. the range of the ball.	
Thinking	**Working**
Write out the information relevant to the horizontal component of the motion. As the ball is hit horizontally, the initial speed gives the horizontal component of the velocity throughout the flight.	Horizontally: v = 25.0 m s^{-1} t = 2.86 s from part (a) s = ?
Select the equation that best fits the information you have.	$v_{av} = \frac{s}{t}$
Substitute values, rearrange and solve for s.	$25.0 = \frac{s}{2.86}$ $s = 25.0 \times 2.86$ $= 71.5$ m Note that the mass of the ball does not affect its motion, as with all objects in projectile motion or in free fall.

c Calculate the velocity of the ball as it lands.	
Thinking	**Working**
Find the horizontal and vertical components of the ball's speed as it lands. Write out the information relevant to both the vertical and horizontal components.	Horizontally: $u = v = 25.0$ m s^{-1} Vertically, with down as positive: $u = 0$ $a = 9.8$ m s^{-2} $s = 40.0$ m $t = 2.86$ s $v = ?$
To find the final vertical speed, use the equation for uniform acceleration that best fits the information you have.	$v = u + at$
Substitute values, rearrange and solve for the variable you are looking for, in this case *v*.	Vertically: $v = 0 + 9.80 \times 2.86$ $= 28.0$ m s^{-1} down
Add the components as vectors.	
Use Pythagoras' theorem to work out the actual speed, *v*, of the ball.	$v = \sqrt{v_h^2 + v_v^2}$ $= \sqrt{25.0^2 + 28.0^2}$ $= \sqrt{1409}$ $= 37.5$ m s^{-1}
Use trigonometry to find the angle, θ.	$\theta = \tan^{-1}\left(\frac{28.0}{25.0}\right)$ $= 48.2°$
Indicate the velocity with a magnitude and direction relative to the horizontal.	The final velocity of the ball is 37.5 m s^{-1} at 48.2° below the horizontal.

Worked example: Try yourself 5.5.1

PROJECTILE LAUNCHED HORIZONTALLY

A golf ball of mass 100 g is hit horizontally from the top of a 30.0 m high cliff with a speed of 20.0 m s^{-1}. Using $g = 9.80$ m s^{-2} and ignoring air resistance, calculate the following values.



a Calculate the time that the ball takes to land.

b Calculate the distance that the ball travels from the base of the cliff, i.e. the range of the ball.

c Calculate the velocity of the ball as it lands.

THE EFFECTS OF AIR RESISTANCE

The interaction between a projectile and the air can have a significant effect on the motion of the projectile, particularly if the projectile has a large surface area and a relatively low mass. If you try to throw an inflated party balloon, it will not travel very far compared to throwing a marble at the same speed.

The size of the **air resistance** or drag force that acts on an object as it moves depends on several factors:

- The speed, v, of the object. The faster an object moves, the greater the drag force becomes.
- The cross-sectional area of the object in its direction of motion. Greater area means greater drag.
- The aerodynamic shape of the object. A more streamlined shape experiences less drag.
- The density of the air. Higher air density means greater drag.

PHYSICSFILE

Aerodynamic design

In the track and field event of javelin, the aerodynamic shape of the javelin was too successful. The javelin was originally designed to reduce the drag force so that the athletes could throw further. This was not a problem until the 1980s, when athletes began to throw so far that runners competing in track events were endangered. The design of the javelin was changed and it was made more snub-nosed to increase drag and reduce distances thrown (this can be seen in Figure 5.5.4). In 1983, the world record was 104.8 m. In 1986 with the modified design, the world record dropped to 85.7 m.

FIGURE 5.5.4 Australia's Kathryn Mitchell in action during the women's javelin final at the London 2012 Olympic Games.

When a pilot drops a supply parcel from a plane, the drag force from the air would act in the opposite direction to the parcel's velocity. If the parcel was dropped on the Moon, where there is no air and hence no air resistance, this would not be a factor and the parcel would continue its horizontal motion and would remain directly below the plane as it fell.

Figure 5.5.5 shows a supply parcel being dropped from a plane moving at a constant velocity. If air resistance is ignored, the parcel falls in the parabolic arc shown by the darker blue curved line in diagram (a). It would continue moving horizontally at the same rate as the plane; that is, as the parcel falls it would stay directly beneath the plane until it hits the ground. The effect of air resistance is also shown by the light-blue curved path in diagram (a). Air resistance (or drag) is a retarding force and it acts in a direction that is opposite to the motion of the projectile. Air resistance makes the parcel fall more slowly, over a shorter path. If air resistance is taken into account, there are now two forces acting, as shown in diagram (b): weight, F_g, and air resistance, F_a. Therefore, the resultant force, F_{net}, that acts on the projectile is not vertically down and nor is its acceleration.

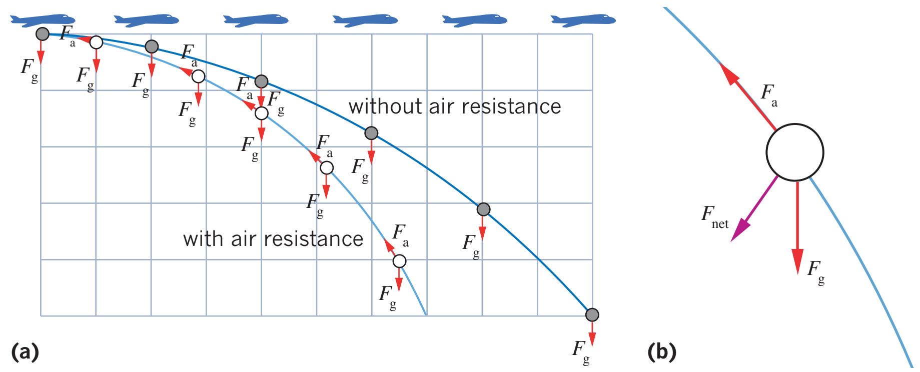


FIGURE 5.5.5 (a) The paths of a supply parcel dropped from a plane with and without air resistance. (b) When air resistance is acting, the net force on the parcel is not vertically down.

5.5 Review

SUMMARY

- If air resistance is ignored, the only force acting on a projectile is its weight, i.e. the force of gravity, F_g. This results in the projectile having a vertical acceleration of 9.8 m s^{-2} down during its flight.
- Projectiles move in parabolic paths that can be analysed by considering the horizontal and vertical components of the motion.
- The following equations of motion for uniform acceleration must be used for the vertical component of the motion:

$v = u + at$

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

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

- The horizontal velocity of a projectile remains constant throughout its flight if air resistance is ignored. Therefore, the following equation for average velocity can be used for this component of the motion:

$v_{av} = \frac{s}{t}$

KEY QUESTIONS

For the following questions, assume that the acceleration due to gravity is 9.8 m s^{-2} and ignore the effects of air resistance unless otherwise stated.

1 A boy throws a stone horizontally at 5 m s^{-1} into a pond. Ignoring air resistance, which statement(s) best describe the stone as it falls towards the pond? More than one option can be correct.

A The only force acting on it is gravity.

B It travels in a circular path.

C There is a driving force acting on it.

D Its speed increases.

2 A marble travelling at 2.0 m s^{-1} rolls off a horizontal bench and takes 0.75 s to reach the floor.

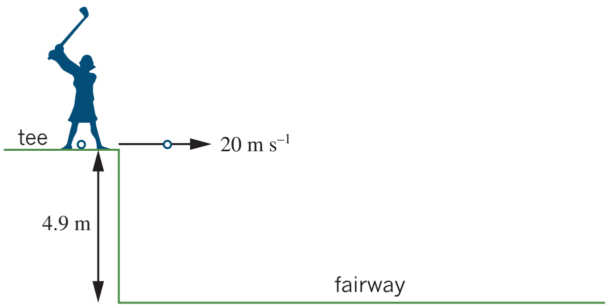
a How far does the marble travel horizontally before landing?

b What is the vertical component of the marble's speed as it lands?

c What is the speed of the marble as it lands?

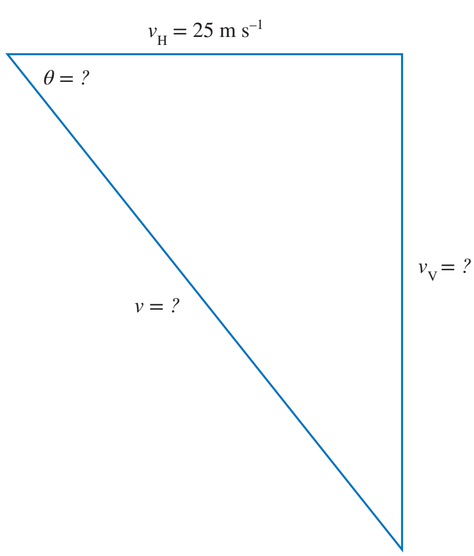
3 A skateboard travelling at 4.0 m s^{-1} rolls off a horizontal bench that is 1.2 m high.
  - **a** How long does the board take to hit the ground?
  - **b** How far does the board land from the base of the bench?
  - **c** What is the magnitude and direction of the acceleration of the board just before it lands?

4 A golfer practising on a range with an elevated tee 4.9 m above the fairway is able to strike a ball so that it leaves the club with a horizontal velocity of 20 m s^{-1}.



  - **a** How long after the ball leaves the club will it land on the fairway?
  - **b** What horizontal distance will the ball travel before landing on the fairway?
  - **c** What is the acceleration of the ball 0.50 s after being hit?
  - **d** Calculate the speed of the ball 0.80 s after it leaves the club.
  - **e** With what speed will the ball hit the ground?

5 A tourist stands on top of a sea cliff that is 80 m high. The tourist throws a rock horizontally at 25 m s^{-1} into the sea.



  - **a** What is the speed of the rock as it reaches the water?
  - **b** At what angle is the rock travelling relative to the horizontal as it reaches the water?

6 In 1971, American astronaut Alan Shepard took a golf club to the Moon and hit a couple of balls. Which of the following answers is/are correct?
  - **A** The balls travelled in straight lines because there is no gravity.
  - **B** The balls travelled in parabolic arcs.
  - **C** The balls travelled much further than if they had been hit in an identical manner on Earth.
  - **D** The balls went into orbit.

7 Joe throws a hockey ball horizontally at 5 m s^{-1}. He then throws a polystyrene ball of identical dimensions at the same speed horizontally. If air resistance is taken into account, which of the balls will travel further? Why?

8 Two identical tennis balls X and Y are hit horizontally from a point 2.0 m above the ground with different initial speeds: ball X has an initial speed of 5.0 m s^{-1}, while ball Y has an initial speed of 10 m s^{-1}.
  - **a** Calculate the time it takes for ball X to strike the ground.
  - **b** Calculate the time it takes for ball Y to strike the ground.
  - **c** How much further than ball X does ball Y travel in the horizontal direction before bouncing?

9 An archer stands on top of a platform that is 20 m high and fires an arrow horizontally at 50 m s^{-1}.
  - **a** What is the speed of the arrow as it reaches the ground?
  - **b** At what angle is the arrow travelling as it reaches the ground, relative to the horizontal?

10 A bowling ball of mass 7.5 kg travelling at 10 m s^{-1} rolls off a horizontal table 1.0 m high.
  - **a** Calculate the ball's horizontal velocity just as it strikes the floor.
  - **b** What is the vertical velocity of the ball as it strikes the floor?
  - **c** Calculate the velocity of the ball as it reaches the floor.
  - **d** What time interval has elapsed between the ball leaving the table and striking the floor?
  - **e** Calculate the horizontal distance travelled by the ball as it falls.
  - **f** Draw a diagram showing the forces acting on the ball as it falls towards the floor.

5.6 Projectiles launched obliquely

The previous section looked at projectiles that were launched horizontally. A more common situation is projectiles that are launched obliquely (at an angle), by being thrown forwards and upwards at the same time. An example of an oblique launch is shooting for a goal in basketball, like in Figure 5.6.1. Once the ball is released, the only forces acting are gravity pulling it down to Earth and air resistance, which retards the ball's motion slightly.

In this section, the principles covering horizontal projectile motion will still apply, as described by Newton's first law.

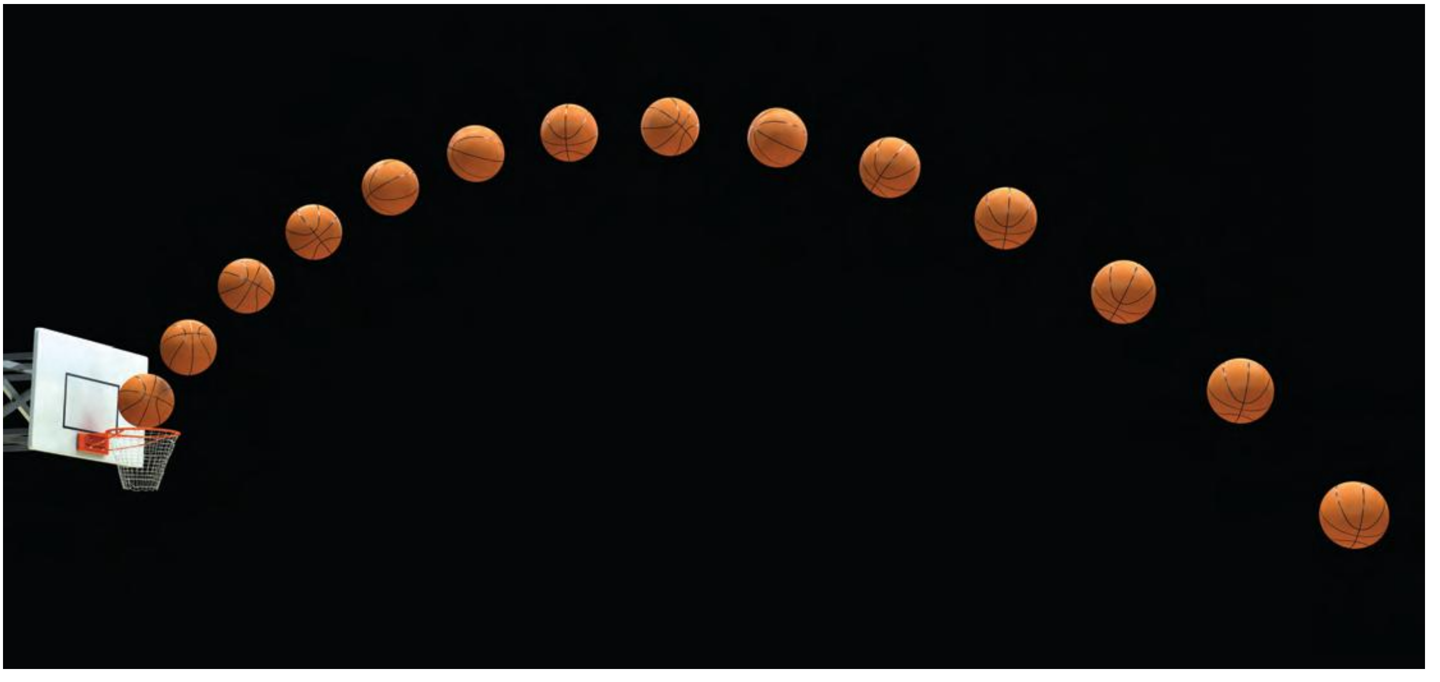


FIGURE 5.6.1 The basketball was thrown up towards the ring, travelling in a smooth parabolic path.

PROJECTILES LAUNCHED AT AN ANGLE

As with projectiles launched horizontally, if drag forces are ignored, the only force that is acting on a projectile that is launched at an angle to the horizontal is gravity, F_g.

Gravity acts vertically downward and so it has no effect on the horizontal motion of a projectile. The vertical and horizontal components of the motion are independent of each other and once again must be treated separately.

In the vertical direction, a projectile accelerates due to the force of gravity, that is, at a rate of 9.8 m s^{-2} downwards. The effect of the force due to gravity is that the vertical component of the projectile's velocity decreases as the projectile rises. It is momentarily zero at the very top of the flight and then it increases again as the projectile descends.

In the horizontal direction a projectile has a uniform velocity since there are no forces acting in this direction (if air resistance is ignored).

Tips for problems involving projectile motion

General rules for solving problems involving projectile motion were given in the previous section—see p180 for a reminder.

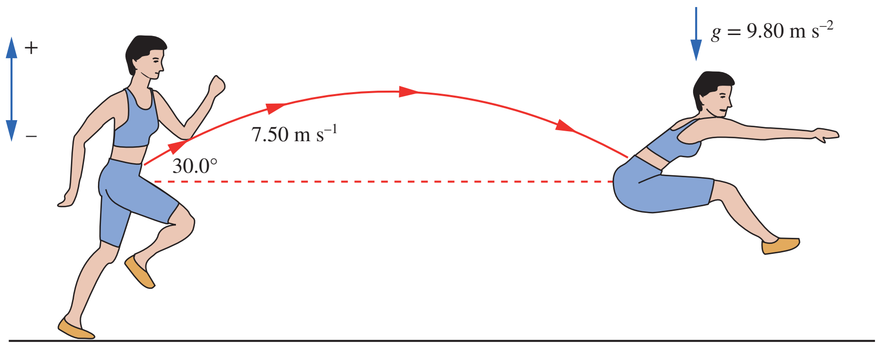
If a projectile is launched at an angle to the horizontal, trigonometry can be used to find the initial horizontal and vertical velocity components. Pythagoras' theorem can be used to determine the actual velocity of the projectile at any point as well as a direction with respect to the horizontal plane.

Worked example 5.6.1 will show you how this is done.

Worked example 5.6.1

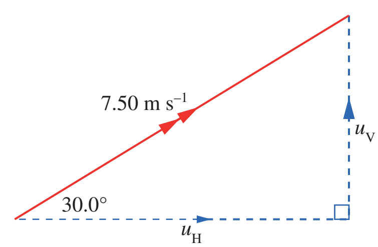
LAUNCH AT AN ANGLE

A 65 kg athlete in a long-jump event leaps with a velocity of 7.50 m s^{-1} at an angle of 30.0° to the horizontal.



For the following questions, treat the athlete as a point mass, ignore air resistance and use g = 9.80 m s^{-2}.

a What is the athlete's velocity at the highest point?

Thinking	Working
First find the horizontal and vertical components of the initial speed.	 Using trigonometry: u_H = 7.50 cos 30.0° = 6.50 m s^{-1} u_V = 7.50 sin 30.0° = 3.75 m s^{-1}
Projectiles that are launched obliquely move horizontally at the highest point. The vertical component of the velocity at this point is therefore zero. The actual velocity is given by the horizontal component of the velocity throughout the motion.	At maximum height: v = 6.50 m s^{-1} horizontally to the right.

b What is the maximum height gained by the athlete's centre of mass during the jump?

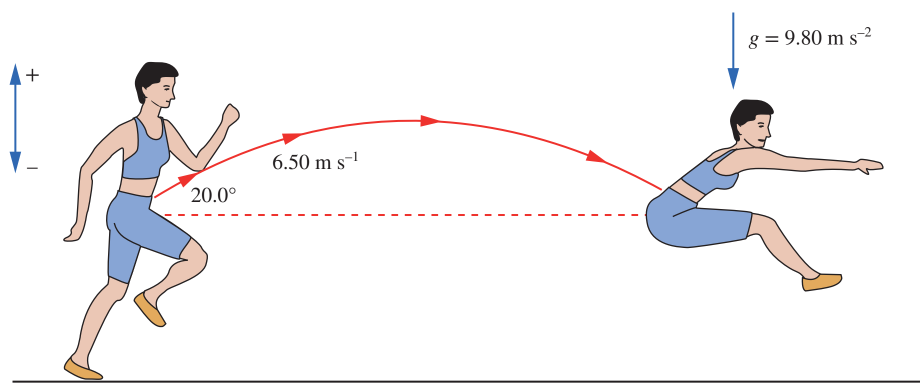
Thinking	Working
To find the maximum height that is gained, you must work with the vertical component. Recall that the vertical component of velocity at the highest point is zero.	Vertically, taking up as positive: u = 3.75 m s^{-1} a = –9.80 m s^{-2} v = 0 s = ?
Substitute these values into an appropriate equation for uniform acceleration.	$v^2 = u^2 + 2as$ $0 = 3.75^2 + 2 \times -9.80 \times s$
Rearrange and solve for s.	$s = \frac{3.75^2}{19.6}$ = 0.717 m

c Assuming a return to the original height, what is the total time the athlete is in the air?	
Thinking	**Working**
As the motion is symmetrical, the time required to complete the motion will be double that taken to reach the maximum height. First, the time it takes to reach the highest point must be found.	Vertically, taking up as positive: $u = 3.75\ \text{m s}^{-1}$ $a = -9.80\ \text{m s}^{-2}$ $v = 0$ $t = ?$
Insert these values into an appropriate equation for uniform acceleration.	$v = u + at$ $0 = 3.75 - 9.80t$
Rearrange the formula and solve for *t*.	$t = \frac{3.75}{9.80}$ $= 0.383\ \text{s}$
The time to complete the motion is double the time it takes to reach the maximum height.	Total time $= 2 \times 0.383$ $= 0.766\ \text{s}$

Worked example: Try yourself 5.6.1

LAUNCH AT AN ANGLE

A 50 kg athlete in a long-jump event leaps with a velocity of $6.50\ \text{m s}^{-1}$ at 20.0° to the horizontal.



For the following questions, treat the athlete as a point mass, ignore air resistance and use $g = 9.80\ \text{m s}^{-2}$.

a What is the athlete's velocity at the highest point?

b What is the maximum height gained by the athlete's centre of mass during the jump?

c Assuming a return to the original height, what is the total time the athlete is in the air?

PHYSICS IN ACTION

Physics of shot putting

In shot-put competitions, there is an advantage in being tall. This enables the release height of the shot to be higher, which means the distance travelled by the shot will be greater. At the London Olympic Games in 2012, the men's event was won by Poland's Tomasz Majewski, with a distance of 21.89 m. Tomasz is 201 cm tall. The gold medal for women was won by Valerie Adams of New Zealand, who threw 20.70 m (see Figure 5.6.2). Valerie is 193 cm tall.

FIGURE 5.6.2 Valerie Adams, of New Zealand, is a tall woman, which helps her to throw the put long distances.

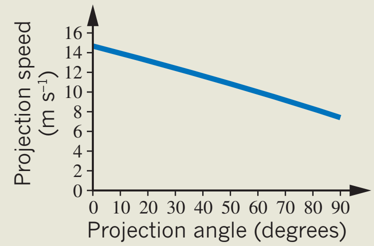


FIGURE 5.6.3 A graph showing how launch speed is greatest with a horizontal launch, and decreases as the launch angle increases.

FIGURE 5.6.4 Tomasz Majewski from Poland won the gold medal for the shot put in London 2012 with a throw of 21.87 m. He would have launched the shot at an angle of around 34°.

When a projectile is launched at an angle to the horizontal, the theoretical launch angle that gives the optimum range is 45°. This applies only where the projectile is launched from zero elevation—that is, when a projectile lands at the same height as it was launched. It is also possible that a projectile lands at a point lower than its launch height. For example with shot putters, the projectile is launched from above the ground. The theoretical launch angle for maximum range in this case is around 43°, depending on the actual release height. In reality, however, shot putters never release at this angle. This is because the speed at which they can launch the shot decreases as the angle gets further from the horizontal. The graph in Figure 5.6.3 shows the relationship between launch speed and launch angle.

The decrease in launch speed with increasing projection angle is caused by two factors:

- When throwing with a high projection angle, the shot putter must expend a greater effort during the delivery phase to overcome the weight of the shot. This reduces the projection speed.
- The structure of the shoulder and arm favours the production of putting force in the horizontal direction more than in the vertical direction.

The optimum projection angle for an athlete is obtained by combining the speed–angle relation for the athlete with the equation for the range of a projectile in free flight. For these reasons, the optimum projection angle for shot putters is actually around 34° (see Figure 5.6.4).

5.6 Review

SUMMARY

- Projectiles move in parabolic paths that can be analysed by considering the horizontal and vertical components of the motion.
- If air resistance is ignored, the only force acting on a projectile is its weight, that is, the force due to gravity, F_g. This results in the projectile having a vertical acceleration of 9.8 m s^{-2} down during its flight.
- The equations for uniform acceleration:

  $v = u + at$

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

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

  must be used for the vertical component.
- The horizontal velocity of a projectile remains constant throughout its flight if air resistance is ignored and $v_{av} = \frac{s}{t}$ is used.
- For objects initially launched at an angle to the horizontal, it is useful to calculate the initial horizontal and vertical velocities using trigonometry.
- At its highest point, the projectile is moving horizontally. Its velocity at this point is given by the horizontal component of its launch velocity. The vertical component of the velocity is zero at this point.

5.6 Review *continued*

KEY QUESTIONS

For the following questions, assume that the acceleration due to gravity is 9.8 m s^{-2} and ignore the effects of air resistance unless otherwise stated.

1 A javelin thrower launches her javelin at 40° above the horizontal. Select the correct statement regarding the javelin at the highest point of its path.

- **A** It has zero acceleration.
- **B** It has its slowest speed.
- **C** There are forwards and downwards forces acting on it.
- **D** There are no forces acting on it since it is in free-fall.

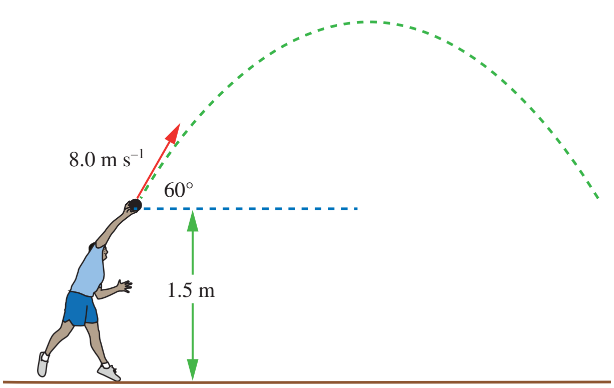
2 A child is holding a garden hose at ground level and the water stream from the hose is travelling at 15 m s^{-1}. Which angle to the horizontal will result in the water stream travelling the greatest horizontal distance through the air?

3 A rugby player kicks for a goal by taking a place kick with the ball at rest on the ground. The ball is kicked at 30° to the horizontal at 20 m s^{-1}. At its highest point, what is the speed of the ball?

4 A basketballer shoots for a goal by launching the ball at 15 m s^{-1} at 25° to the horizontal.

- **a** Calculate the initial horizontal speed of the ball.
- **b** What is the initial vertical speed of the ball?
- **c** What are the magnitude and direction of the acceleration of the ball when it is at its maximum height?
- **d** What is the speed of the ball when it is at its maximum height?

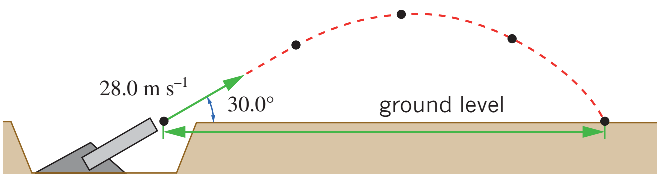
5 In a shot put event a 2.0 kg shot is launched from a height of 1.5 m, with an initial velocity of 8.0 m s^{-1} at an angle of 60° to the horizontal. Answer the questions below about the motion of the shot put.



- **a** What is the initial horizontal speed of the shot?
- **b** What is the initial vertical speed of the shot?
- **c** How long does it take the shot to reach its maximum height?
- **d** What is the maximum height from the ground that is reached by the shot?
- **e** What is the speed of the shot when it reaches its maximum height?

The following information relates to questions 6–11.

A senior physics class conducting a research project on projectile motion constructs a device that can launch a cricket ball. The launching device is designed so that the ball can be launched at ground level with an initial velocity of 28.0 m s^{-1} at an angle of 30.0° to the horizontal.



6 Calculate the horizontal component of the velocity of the ball:

- **a** initially
- **b** after 1.0 s
- **c** after 2.0 s.

7 Calculate the vertical component of the velocity of the ball:

- **a** initially
- **b** after 1.0 s
- **c** after 2.0 s.

8 What is the speed of the cricket ball after 2.00 s?

9 What is the speed of the ball as it lands?

10 What horizontal distance does the ball travel before landing; that is, what is its range?

11 If the effects of air resistance were taken into account, which one of the following statements would be correct?

- **A** The ball would have travelled a greater horizontal distance before striking the ground.
- **B** The ball would have reached a greater maximum height.
- **C** The ball's horizontal velocity would have been continually decreasing.
- **D** The ball's vertical acceleration would have increased.

5.7 Conservation of energy and momentum

Where there are moving objects, there are bound to be collisions. These can range from the interaction of sub-atomic particles, to events on a galactic scale. The Newton's cradle depicted in Figure 5.7.1 provides another example of a collision. The law of conservation of momentum is a powerful tool with which to analyse this motion. The law of conservation of energy is another fundamental principle that can be applied to the interactions between objects, as will be seen in this section.

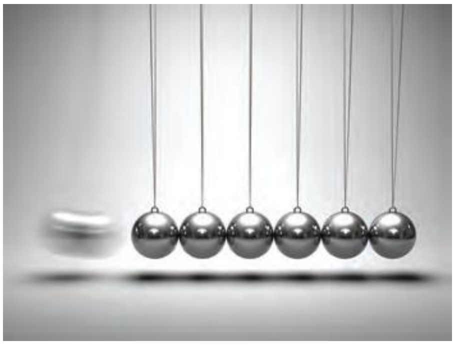


FIGURE 5.7.1 Newton's cradle is a popular illustration of almost perfect conservation of energy and momentum. As the raised sphere loses gravitational potential energy and collides with the other spheres, the sphere at the other end gains kinetic energy. In the collision, the moving sphere transfers its momentum, ejecting a sphere with the same momentum from the end of the row.

CONSERVATION OF MOMENTUM

The product of the mass and velocity of an object is called its **momentum**, and is given by:

$p = mv$

where p is momentum (kg m s^{-1})

m is the mass of the object (kg)

v is the velocity of the object (m s^{-1})

Given that velocity is a vector quantity and momentum is calculated from velocity, it follows that momentum is also a vector quantity.

The **law of conservation of momentum** states that in any collision or interaction between two or more objects in an isolated system, the total momentum of the system will be **conserved** (i.e. it will remain constant). That is, the total initial momentum is equal to the total final momentum:

$\Sigma p_i = \Sigma p_f$

Another way of putting this is that the total change in momentum of the system is zero. This is often found by adding up the change in momentum of all the parts of the system:

$\Sigma \Delta p = 0$

In Physics, a collision refers to a situation when two objects interact and exert action-reaction forces on one another. They do not necessarily have to make physical contact. For instance, two identical charged particles could approach and repel one another, moving off in opposite directions without ever making physical contact.

For the system to be 'isolated', there should only be internal forces acting between the objects, and no interaction with objects outside the system. In reality, perfectly **isolated systems** cannot exist on Earth because of friction and gravity. There are, however, many situations when treating a system as isolated is a useful approximation.

In the rear-end collision between the car and bus examined in Worked example 5.7.1, friction is relatively small compared to the forces exerted by the vehicles on one another. Therefore the vehicles can be treated as an isolated system.

PHYSICSFILE

Discovering the neutrino

Conservation of momentum helped scientists discover the neutrino. In the 1920s it was observed that in a beta decay, a nucleus emitted a beta particle (an electron emitted from a radioactive nucleus). However, when the nucleus recoiled, it was not in the opposite direction to the emitted electron. Thus the momentum of these particles did not comply with the law of conservation of momentum. In 1930 Wolfgang Pauli determined that another particle must also have been emitted in order to conserve the total momentum of the system. This particle, the neutrino, was not detected experimentally until 1956. As you read this, billions of neutrinos are passing through your body and the Earth!

Worked example 5.7.1

CONSERVATION OF MOMENTUM

In a rear-end collision on a freeway, a car of mass 1.0×10^3 kg travelling east at 20 m s^{-1} crashes into the back of a bus of mass 5×10^3 kg travelling east at 8.0 m s^{-1}. Answer the following questions, assuming the car and bus lock together on impact, and ignoring friction.

a Calculate the final common velocity of the vehicles.

Thinking	Working
First assign a direction that will be considered positive.	In this case we will consider vectors directed eastwards to be positive.
Use the law of conservation of momentum.	$\Sigma p_i = \Sigma p_f$ $m_c u_c + m_b u_b = (m_c + m_b)v$ $1.0 \times 10^3 \times 20.0 + 5.0 \times 10^3 \times 8.0 = (1.0 \times 10^3 + 5.0 \times 10^3) \times v$ $60 \times 10^3 = 6.0 \times 10^3 \times v$ $v = 10$ m s^{-1} east

b Calculate the change in momentum of the car.

Thinking	Working
The change in momentum of the car is its final momentum minus its initial momentum.	$\Delta p_c = p_{c(final)} - p_{c(initial)}$ $= m_c (v - u_c)$ $= 1.0 \times 10^3 (10 - 20)$ $= -1.0 \times 10^4$ kg m s^{-1} That is, $\Delta p_c = 1.0 \times 10^4$ kg m s^{-1} west

c Calculate the change in momentum of the bus.

Thinking	Working
The change in momentum of the bus is its final momentum minus its initial momentum.	$\Delta p_b = p_{b(final)} - p_{b(initial)}$ $= m_b (v - u_b)$ $= 5.0 \times 10^3 (10 - 8.0)$ $= 1.0 \times 10^4$ kg m s^{-1} That is, $\Delta p_b = 1.0 \times 10^4$ kg m s^{-1} east

d Verify that the momentum of the system is constant.

Thinking	Working
The total change in momentum is the *vector sum* of the momentum changes of the parts. This is expected to be zero from the conservation of momentum.	$\Delta p_c + \Delta p_b = -1.0 \times 10^4 + 1.0 \times 10^4 = 0$ Therefore the momentum of the system is constant (i.e. conserved) as expected.

Worked example: Try yourself 5.7.1

CONSERVATION OF MOMENTUM

In a head-on collision on a freeway, a car of mass 1.0×10^3 kg travelling east at 20 m s^{-1} crashes into a bus of mass 5.0×10^3 kg travelling west at 8.0 m s^{-1}. Assume the car and bus lock together on impact, and ignore friction.

a Calculate the final common velocity of the vehicles.

b Calculate the change in momentum of the car.

c Calculate the change in momentum of the bus.

d Verify that the momentum of the system is constant.

CONSERVATION OF MOMENTUM FROM NEWTON'S LAWS

The principle of conservation of momentum follows directly from Newton's second and third laws, as illustrated in the tenpin bowling example that follows.

A bowling ball has mass m_b and initial velocity u_b. It collides with the stationary pin of mass m_p.

The ball and pin exert action–reaction forces on one another. In doing so, the velocity of the ball changes and the pin gains a final velocity v_p as shown in Figure 5.7.2.

When the ball and pin collide as shown in the middle diagram of Figure 5.7.2, they exert action–reaction forces on one another, and according to Newton's third law:

$F_{\text{on b by p}} = -F_{\text{on p by b}}$

The forces cause the ball to decelerate and the pin to accelerate, so from Newton's second law, $F = ma$:

$m_b a_b = -m_p a_p$

The ball and pin are in contact for time Δt, and so rewriting acceleration in terms of velocity gives:

$m_b \frac{(v_b - u_b)}{\Delta t} = -m_p \frac{(v_p - u_p)}{\Delta t}$

The times are the same and so they cancel, hence:

$m_b (v_b - u_b) = -m_p (v_p - u_p)$

Expanding and rearranging gives:

$m_b v_b + m_p v_p = m_b u_b + m_p u_p$

Swap sides as the convention is to write the initial momenta on the left and the final momenta on the right. This gives:

$m_b u_b + m_p u_p = m_b v_b + m_p v_p$

The left-hand side of the equation above describes the initial momentum of the system while the right-hand side represents the final momentum of the system.

Thus the application of Newton's second and third laws has produced the conservation of momentum equation:

$\Sigma p_i = \Sigma p_f$

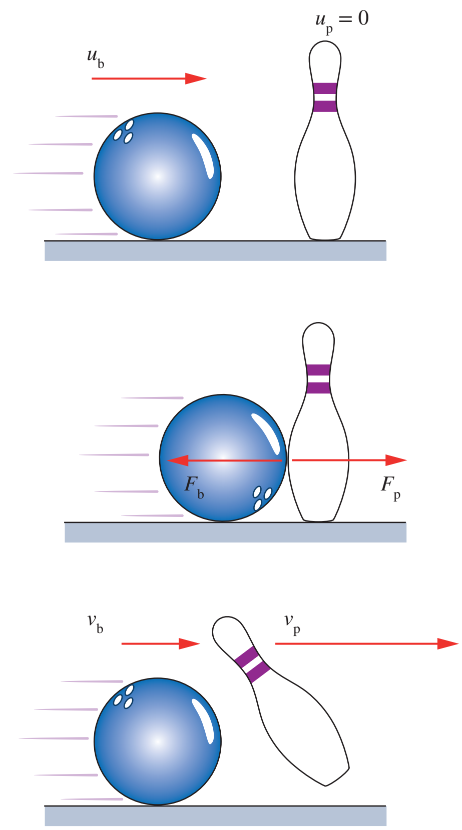


FIGURE 5.7.2 When a bowling ball collides with a pin, they exert equal but opposite forces on each other. These forces cause the ball to lose some momentum and the pin to gain an equal amount of momentum.

PHYSICSFILE

Not so strongman

Traditionally, circus strongmen would often perform a feat where they place a large rock on their chest, then allow another person to smash the rock with a sledgehammer. This might seem at first to be an act of extreme strength and daring. However, a quick analysis using the principle of conservation of momentum will show otherwise. Assume that the rock has a mass of 12 kg and that the sledgehammer of mass 3 kg strikes it at 5 m s^{-1}. Using conservation of momentum, we can show that the rock and sledgehammer will move together at just 1 m s^{-1} after the impact:

$m_1 u_1 + m_2 u_2 = (m_1 + m_2)v$

$3 \times 5 + 12 \times 0 = (3 + 12) \times v$

$15 = 15v$

$v = 1$ m s^{-1}

The large mass of the rock has dictated that the final common speed is too low to hurt the strongman. A more daring feat would be to use the sledgehammer to smash a pebble!

CONSERVATION OF ENERGY

Energy comes in many forms, such as heat, light, sound, chemical energy and electrical energy. It is a scalar quantity and is measured in joules (J). Energy is also associated with the motion and position of an object, and this energy is called the **mechanical energy** of the object. In the motion problems of this chapter, moving objects have been described as having **kinetic energy**. An object can also have stored or **gravitational potential energy** because of its position. For instance, a building crane lifting a steel beam several stories has given the beam gravitational potential energy that could be disastrously converted to kinetic energy if the lifting chain were to break and the beam were to accelerate under gravity.

The changing of gravitational potential energy to kinetic energy is an illustration of the **law of conservation of energy**, a fundamental principle of nature. The law of conservation of energy states that energy is not created or destroyed, but can only change from one form to another, or in other words, **transform**. As the gravitational potential energy of the beam decreases, its kinetic energy increases. The total amount of mechanical energy remains constant.

While energy is not ever destroyed, it may be dissipated in forms that are not easily recoverable. For instance, the kinetic energy of a vehicle is reduced as it encounters friction, causing heating of the tyres, or in the deformation of the bodywork should it collide with another object. The mechanical energy before and after a collision is only the same under ideal conditions, but in many cases it is a useful approximation.

Gravitational potential energy

The energy that an object gains due to its position in a gravitational field is called the gravitational potential energy. Close to the surface of the Earth this is given by:

$E_g = mg\Delta h$

where E_g is the gravitational potential energy gained or lost (J)

m is the mass of the object (kg)

g is the gravitational field strength (N kg^{-1})

Δh is the change in height of the object (m)

Kinetic energy

The energy of motion depends on the mass of the object and its speed:

$E_k = \frac{1}{2}mv^2$

where E_k is the kinetic energy of the object in motion (J)

m is the mass of the object (kg)

v is the speed of the object (m s^{-1})

Problems combining gravitational potential and kinetic energy

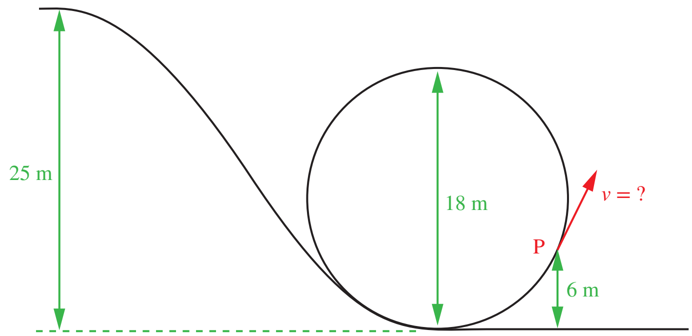
Energy is a scalar quantity; hence it can be easier to work with compared with vector quantities. Therefore it is worth analysing a problem to see if calculations involving energy are possible before using techniques involving forces and other vectors.

The speed of an object can be determined from its kinetic energy, even if its mass is unknown. This is best demonstrated by an example.

Worked example 5.7.2

CONSERVATION OF ENERGY

Consider a rollercoaster with a lift hill of height 25.0 m and a loop height of 18.0 m as shown in the figure below. At the top of the lift hill, the rollercoaster car has zero velocity; then it begins to roll down the hill. Calculate the speed of the car at point P on the loop, when the car is 6 m above the ground. Assume friction is negligible.

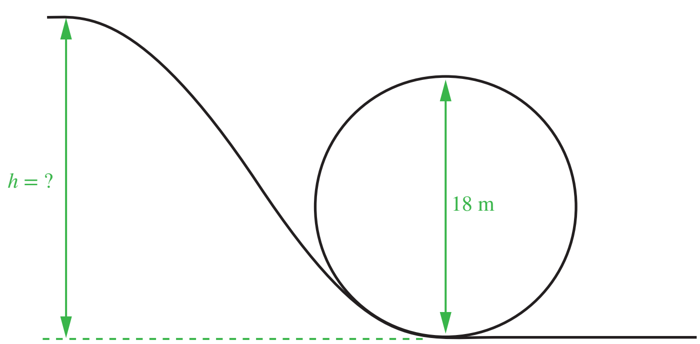


Thinking	Working
Find the total mechanical energy, E_m, of the cart from the gravitational potential energy at the lift hill at the start of the ride.	$E_g = mg\Delta h$ $= (m \times 9.8 \times 25)$ $= 245m$ J $E_k = 0$ so $E_m = 245m$ J
At point P, E_m consists of gravitational potential energy and kinetic energy. Write a statement for E_m at point P.	$E_m = (m \times 9.8 \times 6) + \frac{1}{2}mv^2$ $= 58.8m + \frac{1}{2}mv^2$
Equate the two statements for E_m.	$245m = 58.8m + \frac{1}{2}mv^2$
Cancel *m* from both sides and rearrange.	$245 = 58.8 + \frac{1}{2}v^2$ $(245 - 58.8) \times 2 = v^2$
Solve for the speed.	$v = 19$ m s^{-1}

Worked example: Try yourself 5.7.2

CONSERVATION OF ENERGY

Use the law of conservation of energy to determine the height of the lift hill required to ensure that the speed of the car at the top of the 18.0 m loop is 25.0 m s^{-1}. Assume that the velocity of the car at the top of the lift hill is zero; then it begins to roll down the hill.



5.7 Review

SUMMARY

- The momentum of an object is the product of its mass and its velocity: $p = mv$. Momentum is measured in kg m s^{-1} or N s.
- The total momentum of an isolated system is conserved. Therefore, the vector sum of momentum of the parts of a system before a collision is equal to the total momentum after the collision: $\Sigma p_i = \Sigma p_f$.
- In a simple collision between two objects of mass m_1 and m_2 this becomes:
  $m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2$
- Energy is not created or destroyed, but merely transformed. This is called the law of conservation of energy.
- Energy is a scalar quantity and is measured in joules (J).
- The energy an object has because of its position in a gravitational field is called gravitational potential energy. Close to the surface of the Earth this may be found from:
  $E_g = mg\Delta h$
- The energy of movement is called kinetic energy and is calculated using:
  $E_k = \frac{1}{2}mv^2$
- Because it is simpler to work with scalars, it is often helpful to solve motion problems using energy considerations.

KEY QUESTIONS

1 Two billiard balls are rolled very carefully from opposite ends of a pool table. They collide in the middle and both balls come to rest. Has momentum been conserved in this system? Explain your answer.

2 A 175 kg sumo wrestler running east with a speed of 3.5 m s^{-1} crashes into an opponent of mass 100 kg running in the opposite direction at 5.0 m s^{-1}. The two wrestlers collide while in mid-air, and remain locked together after the collision. Calculate their final velocity.

3 A 110 kg ice hockey player travelling at 15 m s^{-1} collides with a player of mass 90 kg who is travelling at 5.0 m s^{-1} towards him. The two players remain locked together after the collision as they slide across the ice. Ignoring friction, find their joint speed.

The following information applies to questions 4 and 5.

A sports car of mass 1.0×10^3 kg travelling east at 36 km h^{-1} approaches a station wagon of mass 2.0×10^3 kg moving west at 18 km h^{-1}.

4 **a** Calculate the momentum of the sports car.
b Calculate the momentum of the station wagon.
c Determine the total momentum of these vehicles.

5 These two vehicles now collide head-on on an icy stretch of road where there is negligible friction. The vehicles remain locked together after the collision.
a Calculate the common velocity of the two vehicles after the collision.
b Where has the initial momentum of the vehicles gone?
c Determine the change in momentum of the sports car.
d Determine the change in momentum of the station wagon.

6 A 200 g snooker ball travelling with initial velocity 9.0 m s^{-1} to the right collides with a stationary ball of mass 100 g. If the final velocity of the 200 g ball is 3.0 m s^{-1} to the right, calculate the velocity of the 100 g ball after the collision.

7 A 1000 kg cannon mounted on wheels fires a 10.0 kg shell with a horizontal speed of 500 m s^{-1}. Assuming that friction is negligible, calculate the recoil velocity of the cannon.

8 An arrow of mass 100 g is fired with an initial horizontal velocity of 40 m s^{-1} to the right at an apple of mass 80 g that is initially at rest on a horizontal surface. When the arrow strikes the apple, the two objects stick together. What is the common velocity of the arrow and apple after the impact?

9 A 40.0 g bullet is fired into the air with a muzzle velocity of 370 m s^{-1}. Calculate the kinetic energy of the bullet as it leaves the firearm.

10 Calculate the gravitational potential energy that a 40.0 g bullet has when it has travelled 1000 m up into the air after having been fired from a gun.

11 A boy throws a 157 g cricket ball up into the air. It leaves his hand at a speed of 20.5 m s^{-1}.
a Calculate the kinetic energy of the ball as it leaves the boy's hand.
b What is the gravitational potential energy at the top of its flight, if air resistance is ignored?
c Calculate the maximum height reached, if air resistance is ignored.

Chapter review

05

KEY TERMS

air resistance
banked track
centripetal acceleration
centripetal force
conserved
design speed
free fall
frequency
gravitational potential energy
inclined plane
isolated system
kinetic energy
law of conservation of energy
law of conservation of momentum
magnitude
mechanical energy
momentum
normal reaction force
period
projectile
tangential
transform

1 Diana rolls a bowling ball down a smooth straight ramp. Choose the option below that best describes the way the ball will travel.
  A with constant speed
  B with constant acceleration
  C with decreasing speed
  D with increasing acceleration

2 Two blocks are joined by a string that passes over a frictionless pulley. One of the blocks is placed on a frictionless table and the other is free to fall. The block on the table has a mass of 5 kg and the falling block has a mass of 10 kg.
  a At what rate do the blocks accelerate?
  b What is the magnitude of the tension in the string?

3 A 1000 kg car tows a 200 kg trailer along a level surface with an acceleration of 2.5 m s^{-2}. The frictional drag on the car is 800 N and the drag on the trailer is 700 N. Calculate the thrust provided by the car engine to give this acceleration.

4 A bowling ball is rolling down a smooth track that is inclined at 30° to the horizontal.
  a What is the magnitude of the acceleration of the ball?
  b How does the magnitude of the normal force that acts on the ball compare to its weight?

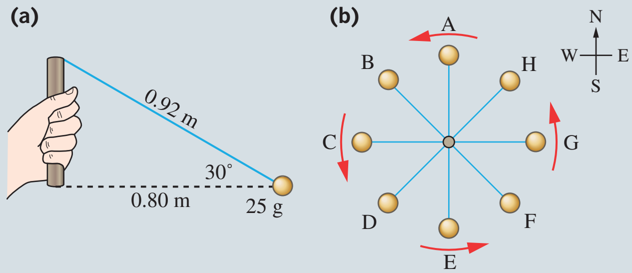
5 A marble is rolled from rest down a smooth slide that is 2.5 m long. The slide is inclined at an angle of 30° to the horizontal.
  a Calculate the acceleration of the marble.
  b What is the speed of the marble as it reaches the end of the slide?

6 Marshall has a mass of 57 kg and he is riding his 3.0 kg skateboard down a 5.0 m long ramp that is inclined at an angle of 65° to the horizontal. Ignore friction when answering questions **a** to **d**.
  a Calculate the magnitude of the normal force acting on Marshall and his skateboard.
  b What is the acceleration of Marshall as he travels down the ramp?
  c What is the net force acting on Marshall and his board when no friction acts?
  d If Marshall's initial speed is zero at the top of the ramp, calculate his final speed as he reaches the bottom of the ramp.
  e Marshall now stands halfway up the incline while holding his board in his hands. Friction now acts on Marshall. Calculate the frictional force acting on Marshall while he is standing stationary on the ramp.

7 The highest waterslide in the world is in the United States of America. It is 50.0 m tall and is inclined at an angle of 70° to the horizontal. It is known that riders reach a speed of 100 km h^{-1} on this slide. Do not assume friction is negligible.
  a For a 70.0 kg teenager using the slide, calculate the net force on the teenager as he slides.
  b For the same teenager, calculate the magnitude of the average frictional force opposing the motion.
  c If the friction acts on the teenager to slow him down, what is the reaction force to this?
  d What is the reaction force to the teenager's weight force?

The following information applies to questions 8 and 9.

During a high-school physics experiment, a copper ball of mass 25.0 g was attached to a very light piece of steel wire 0.920 m long and whirled in a circle at 30.0° to the horizontal, as shown in diagram (a). The ball moves in a circular path of radius 0.800 m with a period of 1.36 s. The top view of the resulting motion of the ball is shown in diagram (b).



8 **a** Calculate the orbital speed of the ball.
b What is the centripetal acceleration of the ball?
c What is the magnitude of the centripetal force acting on the ball?

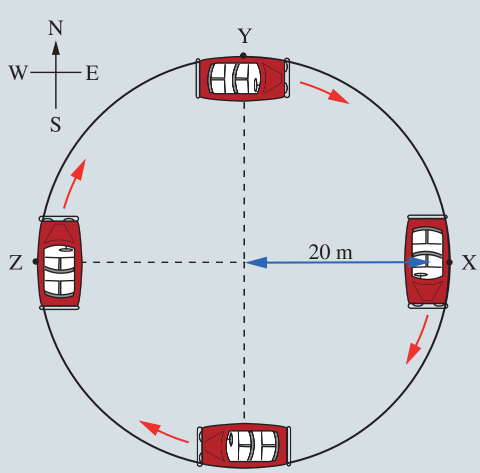
9 **a** Draw a diagram similar to diagram (a) that shows all the forces acting on the ball at this time.
b What is the magnitude of the tension in the wire?

10 A radio-controlled car is travelling in a circular path of radius 10 m at a constant speed of 5.0 m s^{-1}.
a What is the acceleration of the toy car?
b What force is creating the circular motion of the car?

11 The Moon orbits the Earth once in 27.3 days in a circular orbit of radius 3.84×10^8 m.
a Calculate the orbital speed of the Moon.
b Calculate the net force keeping the Moon in orbit if the mass of the Moon is 7.36×10^{22} kg.

12 A geostationary communications satellite is at an altitude of 3.60×10^4 m. The Earth has an average radius 6.37×10^6 m and a period of rotation of 23 hours, 56 minutes and 5 seconds. Calculate the centripetal acceleration of the satellite.

13 A car of mass 1500 kg is driven at constant speed of 10 m s^{-1} around a level, circular roundabout. The centre of mass of the car is always 20 m from the centre of the track.



a What is the velocity of the car at point X?
b What is the speed of the car at point Y?
c What is the period of revolution for this car?
d What is the acceleration of the car at point X?
e Determine the size and direction of the unbalanced frictional force acting on the tyres at point X.

14 A proton moves into a region of uniform magnetic field 0.250 T directed perpendicular to the velocity vector. If it travels into the field at 3.50×10^6 m s^{-1}, calculate the radius of curvature of its path. Note that: $m_p = 1.67 \times 10^{-27}$ kg and $q = 1.60 \times 10^{-19}$ C.

15 A track cyclist is riding at high speed on the steeply banked section of a velodrome ($\theta = 37°$). Which statement describes the size of the normal force acting on the cyclist at this point?
A greater than the weight of the cyclist
B zero
C less than the weight of the cyclist
D equal to the weight of the cyclist

16 A cycling track has a turn that is banked at 40° to the horizontal. The radius of the track at this point is 30 m. Determine the speed at which a cyclist of mass 60 kg would experience no sideways force on their bike as they ride this section of track.

17 The Ferris wheel at an amusement park has an arm radius of 10 m and its compartments move with a constant speed of 5.0 m s^{-1}.
a Calculate the normal force that a 50 kg boy would experience from the seat when at the:
i top of the ride
ii bottom of the ride.
b After getting off the ride, the boy remarks to a friend that he felt lighter than usual at the top of the ride. Which option explains why he might feel lighter at the top of the ride?
A He lost weight during the ride.
B The strength of the gravitational field was weaker at the top of the ride.
C The normal force there was larger than the gravitational force.
D The normal force there was smaller than the gravitational force.

18 As part of a demonstration, a teacher swings a bucket half-filled with water in a vertical circle at high speed. No water spills from the bucket even when it passes the overhead position. Discuss the forces acting on the water when the bucket is directly overhead and indicate their directions. Indicate the direction and relative size of the water's acceleration as it passes this position.

19 A toy car is moving at 2.5 m s^{-1} as it rolls off a horizontal table. The car takes 1.0 s to reach the floor.

a How far does the car land from the foot of the table?

b What are the magnitude and direction of acceleration when the car is halfway to the floor?

20 A bowling ball of mass 7.5 kg travelling at 10 m s^{-1} rolls off a horizontal table that is 0.97 m high. Answer the questions below about the motion of the bowling ball.

a What is the horizontal speed of the ball as it strikes the floor?

b What is the vertical speed of the ball as it strikes the floor?

c Calculate the speed of the ball as it reaches the floor.

21 In a tennis match, a tennis ball is hit from a height of 1.2 m with an initial velocity of 16 m s^{-1} at an angle of 50° to the horizontal. Ignore drag forces for the following questions.

a What is the initial horizontal speed of the ball?

b What is the initial vertical speed of the ball?

c What is the maximum height that the ball reaches above the court surface?

22 An orange of mass 100 g is tossed horizontally at 6.0 m s^{-1} from a height of 2.0 m. Ignore air resistance and use g = 9.8 m s^{-2} when answering these questions.

a What is the initial kinetic energy of the orange?

b Calculate the initial potential energy of the orange.

c What is the speed of the orange as it lands?

The following information applies to questions 23–26.

A 50 kg boy stands on a 200 kg sled that is at rest on a frozen pond. The boy jumps off the sled with a velocity of 4.0 m s^{-1} east.

After the boy has jumped off, he turns around and skates after the sled, jumping on with a horizontal velocity of 4.4 m s^{-1} west.

23 a What is the total momentum of the boy and the sled before he jumps off?

b What is the momentum of the boy after he jumps?

c What is the momentum of the sled after he has jumped?

24 Assuming the pond surface is frictionless, calculate the velocity of the sled just before he jumps on?

25 What is the speed of the boy once he is on the sled?

26 As the boy jumps on the sled, what change in momentum is experienced by the:

a sled

b boy?


Santander
V-Power
Santander
telcel
MAHLE
ups
Santander
WEICHAI
HUBLOT
brembo

CHAPTER 06

Special relativity

Galileo and Newton laid the foundations of the 'clockwork universe', a mechanical picture of the world which has underpinned most modern world views. Einstein, along with others such as Bohr and Heisenberg, presented a much richer and more mysterious universe, one that challenges people to think beyond the mechanical picture they so often take for granted.

In this chapter, you will explore the concepts of classical physics, as described by Galileo and Newton, and the evidence that pointed towards the need for some different thinking. Einstein's special relativity is presented as a solution to the problem of classical physics at speeds approaching the speed of light.

Key knowledge

By the end of this chapter, you will have covered material from the study of relativity and will be able to:

- describe Einstein's two postulates for his theory of special relativity that:
  - the laws of physics are the same in all inertial (non-accelerated) frames of reference
  - the speed of light has a constant value for all observers regardless of their motion or the motion of the source
- compare Einstein's theory of special relativity with the principles of classical physics
- describe proper time (t_0) as the time interval between two events in a reference frame where the two events occur at the same point in space
- describe proper length (L_0) as the length that is measured in the frame of reference in which objects are at rest
- model mathematically time dilation and length contraction at speeds approaching c using the equations: $t = t_0\gamma$ and $L = \frac{L_0}{\gamma}$ where $\gamma = \left(1-\frac{v^2}{c^2}\right)^{-\frac{1}{2}}$
- explain why muons can reach Earth even though their half-lives would suggest that they should decay in the outer atmosphere.

6.1 Einstein's theory of special relativity

Galileo and Newton developed theories of motion. These theories allowed the relative motion of low-speed objects to be modelled mathematically. This section presents the observations that challenged Galilean relativity and Newtonian physics, and explains the key principles that led to the new physics described by Albert Einstein (see Figure 6.1.1), known as the theory of special relativity.

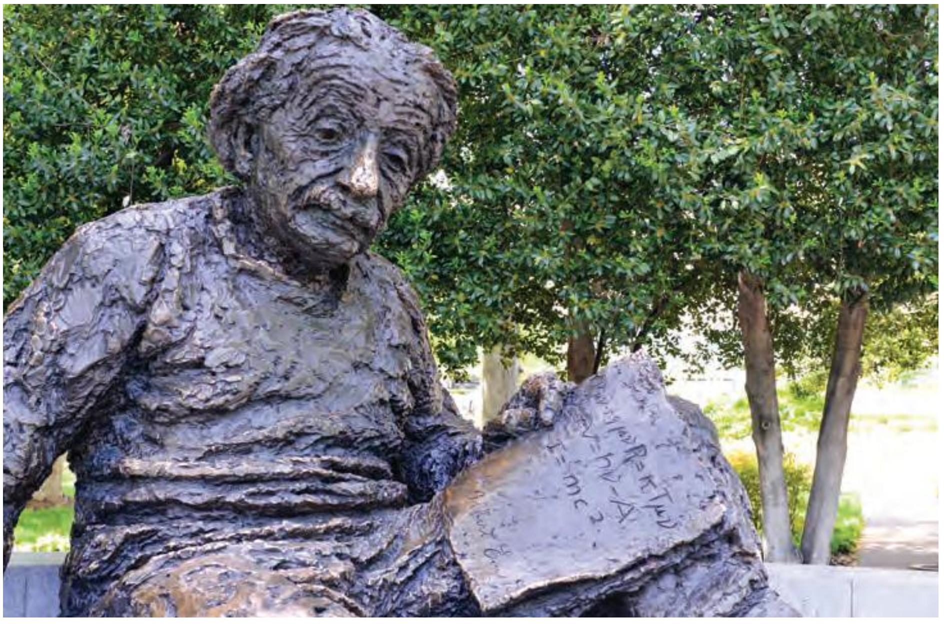


FIGURE 6.1.1 Albert Einstein statue in Washington, D.C.

EINSTEIN'S BRILLIANT THEORY

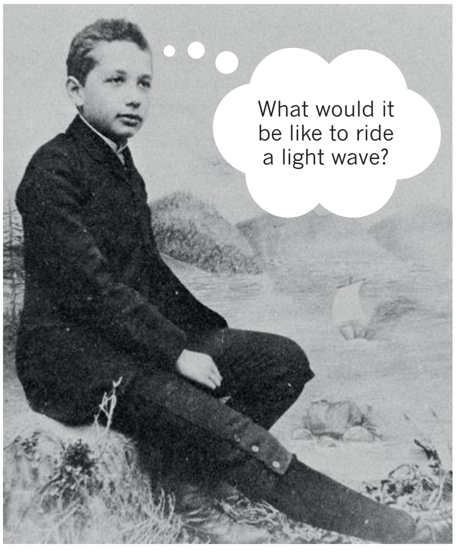


FIGURE 6.1.2 Einstein as a teenager.

When Albert Einstein was just 5 years old, his father gave him a compass. He was fascinated by the fact that it was responding to some invisible field that enveloped the Earth. His curiosity was aroused and, fortunately for physics, he never lost it. In his teens he turned his attention to the question of light (see Figure 6.1.2).

Perhaps it was lucky that in his early twenties Einstein was not part of the physics 'establishment'. He was working as a patent clerk in the Swiss Patent Office. It was an interesting enough job, but it left him time to think about electromagnetic waves (light) and their relationship to the Galilean principle of relativity.

Galileo was particularly interested in relative motion. One of his famous experiments involved the dropping of a cannon ball from the top of the mast of a moving ship. Galileo found that the motion of the cannon ball was not affected by the motion of the ship; the cannon ball landed next to the base of the mast. His principle of relativity was that you cannot tell if you are moving or not without looking outside of your own **frame of reference**.

Based on the work of Galileo, Isaac Newton established detailed models for the motion of objects such as planets, moons and comets, even falling oranges. According to his equations, the velocity of objects can be calculated relative to any frame of reference as long as the velocity of the frame of reference is known. The Newtonian principle that the velocities of objects and frames of reference can be added together to determine the velocity of the object in another frame of reference is common throughout his equations and laws.

Consider an object moving in a frame of reference, A. This frame of reference is moving in another frame of reference, B. The velocity of the object in B is given by:

$$v_{\text{object in B}} = v_{\text{object in A}} + v_{\text{A in B}}$$

A practical example of this could be when a person runs forward on a train. Here, the train is frame of reference A and the track along which the train moves is frame B. Imagine that the person runs at 5 m s^{-1} forwards, while the train travels at a velocity of 20 m s^{-1} forwards. The velocity of the person relative to B, the track, is:

$$v_{\text{person along track}} = v_{\text{person in train}} + v_{\text{train along track}}$$
$$= 5 + 20$$
$$= 25 \text{ m s}^{-1}$$

That is, the person is moving with a velocity of 25 m s^{-1} forwards when measured against the track.

Einstein was a typical theoretician; the only significant experiments he ever did were thought experiments, or ***Gedanken*** experiments, as they are called in German. Many of his *Gedanken* experiments involved thinking of situations that involved two frames of reference moving with a steady relative velocity, in which the principles of Galilean relativity applied. Newton had referred to these as **inertial frames of reference**, as the law of inertia applied within them.

Einstein and Galilean relativity

Einstein decided that the elegance of the principle of Galilean relativity was such that it simply had to be true. Nature did not appear to have a special frame of reference, and Einstein could see no reason to believe that there was one waiting to be discovered. In other words, there is no such thing as an absolute velocity. It is not possible to have a velocity relative to space itself, only to other objects within space. So the velocity of any object can always be stated as relative to some other object. In the case of the person running on the train, their velocity can be stated as either 5 m s^{-1} relative to the train or 25 m s^{-1} relative to the track.

Einstein expanded the Galilean principle to state that all inertial frames of reference must be equally valid, and that the laws of physics must apply equally in any frame of reference that is moving at a constant velocity. So there is no physics experiment you can do that is entirely within a frame of reference that will tell you that you are moving. In other words, as you speed along in your *Gedanken* train with the blinds down, you cannot measure your speed. You can tell if you are accelerating easily enough: just hang a pendulum from the ceiling. However, the pendulum will hang straight down whether you are travelling steadily at 100 km h^{-1} or are stopped at the station. Consider Figure 6.1.3(a) and (b). There is no way of telling which of the trains is stationary relative to the ground, or which is moving at a constant velocity.

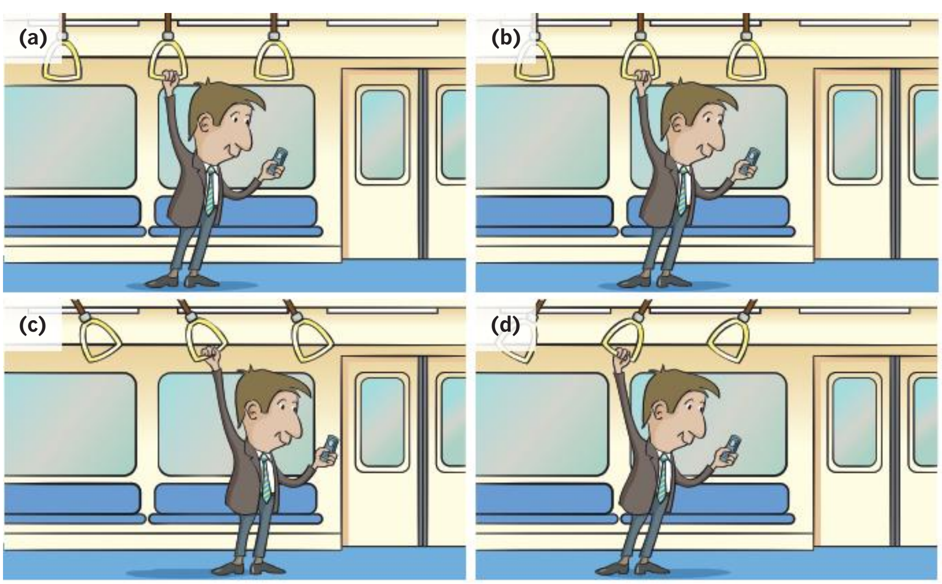


FIGURE 6.1.3 There is no observation or experiment that shows the difference between two inertial frames of reference (a) and (b). In one of the situations illustrated, the train is stationary, and in the other it is moving smoothly at 100 km h^{-1}. There is no observation that will tell which one is which. In (c) and (d), the motion of the handles hanging from the ceiling of the train indicate that these trains are not moving at a constant speed.

Einstein decided that the relativity principle could not be abandoned. Recall that Einstein was, at the time, thinking about the relationship between light and relativity. Whatever the explanation for the strange behaviour of light, it could not be based on a flaw in the principle of Galilean relativity.

Einstein's fascination with the nature of light had led him to a deep understanding of Maxwell's work on the electromagnetic nature of light waves. He was convinced of the elegance of Maxwell's equations and their prediction of a constant speed of light. Most physicists believed that the constant speed predicted by Maxwell's equations referred to the speed of light relative to a **medium** (a substance it travelled through). It was thought that the speed predicted would be the speed in the medium in which light travelled, and the measured speed would have to be adjusted for one's own speed through that medium.

As light travelled through the vacuum of space between the Sun and Earth, clearly the medium was no ordinary material. Physicists gave it the name **aether**, as it was an 'ethereal' substance. It was thought, following Maxwell's work, that the aether must be some sort of massless, rigid medium that 'carried' electric and magnetic fields.

This was a real problem for Einstein. A speed of light that is fixed in the aether and which depended on the velocity of an inertial frame in the aether would be in direct conflict with the principle of Galilean relativity, which Einstein was reluctant to abandon.

Resolving the problem of the aether

As in any conflict, the resolution is usually found by people who are prepared to look at it in new ways. This was the essence of Einstein's genius. Instead of looking for faults in what appeared to be two perfectly good principles of physics, he decided to see what happened if they were both accepted, despite the apparent contradiction.

So Einstein swept away the problem of the aether, saying that it was simply unnecessary. It had been invented only to be a medium for light waves, and no one had found any evidence for its existence (refer to the Extension box on page 205). Electromagnetic waves, he said, could apparently travel through empty space without a medium. Doing away with the aether, however, did not solve the basic conflict between the absolute speed of light and the principle of relativity.

EINSTEIN'S THEORY OF SPECIAL RELATIVITY

Though Einstein accepted both Galileo's and Maxwell's theories despite the apparent contradiction, this still left the question: How could two observers travelling at different speeds see the same light beam travelling at the same speed? The answer, Einstein said, was in the very nature of space and time.

In 1905 he sent a paper to the respected physics journal *Annalen der Physik* entitled 'On the electrodynamics of moving bodies'. In this paper he put forward two simple **postulates** (statements assumed to be true) and followed them to their logical conclusion. It was this conclusion that was so astounding.

Einstein's two postulates:

- The laws of physics are the same in all inertial (non-accelerated) frames of reference.
- The speed of light has a constant value for all observers regardless of their motion or the motion of the source.

(The first postulate means that there is no preferred frame of reference and so is sometimes stated as: no law of physics can identify a state of absolute rest.)

EXTENSION

The Michelson–Morley experiment

The existence of an aether appeared to be a serious blow for the principle of relativity. It seemed that there may be after all, a frame of reference attached to space itself. If this was the case, there was the possibility of an absolute zero velocity.

Scientists needed to test the idea of electromagnetic waves moving through the aether. Since the Earth is in orbit around the Sun, an aether wind should be blowing past the Earth. This suggested to American physicist Albert Michelson that it should be possible to measure the speed at which the Earth was moving through the aether by measuring the small changes in the speed of light as the Earth changed its direction of travel. For example, if the light was travelling in the same direction as the Earth, through the aether, the apparent speed should be slower than usual at $c - v$ (see Figure 6.1.4). It would be as if the light was travelling against an aether 'wind' created by the motion of the Earth through it. If the light was travelling against the Earth's motion, the apparent speed should be faster as it would be travelling with the 'wind' at $c + v$ (see Figure 6.1.4). The differences would be tiny, less than 0.01%, but Michelson was confident that he could measure them.

In the 1880s Michelson, and his collaborator Edward Morley, set up a device known as an interferometer. The device cannot measure the speed of light but it can detect changes in the speed of light that might have been due to the aether wind. In fact, it was used to attempt to measure the very small differences in the time taken for light to travel in two mutually perpendicular directions. They were able to rotate the whole apparatus and hoped to detect the small difference that should result from the fact that one of the directions was to be the same as that in which the Earth was travelling and the other at right angles. However, they found no difference. Perhaps, then, the Earth at that time was stationary with respect to the aether? Six months later, however, when the Earth would have to be travelling in the opposite direction relative to the aether, there was still no difference in the measured speeds! Other people performed similar experiments, virtually always with the same null result. Whatever direction the Earth was moving it seemed to be at rest in the aether. Or perhaps there was no aether at all.

While Michelson and Morley's results were consistent with Maxwell's prediction that the speed of light would always appear to be the same for any observer, the apparent absurdity of such a situation led most physicists to believe that some flaw in the theory behind the experiment, or in its implementation, would soon be discovered. Einstein, however, wondered about the consequences of actually accepting their prediction about the speed of light but at the same time holding on to the relativity principle.

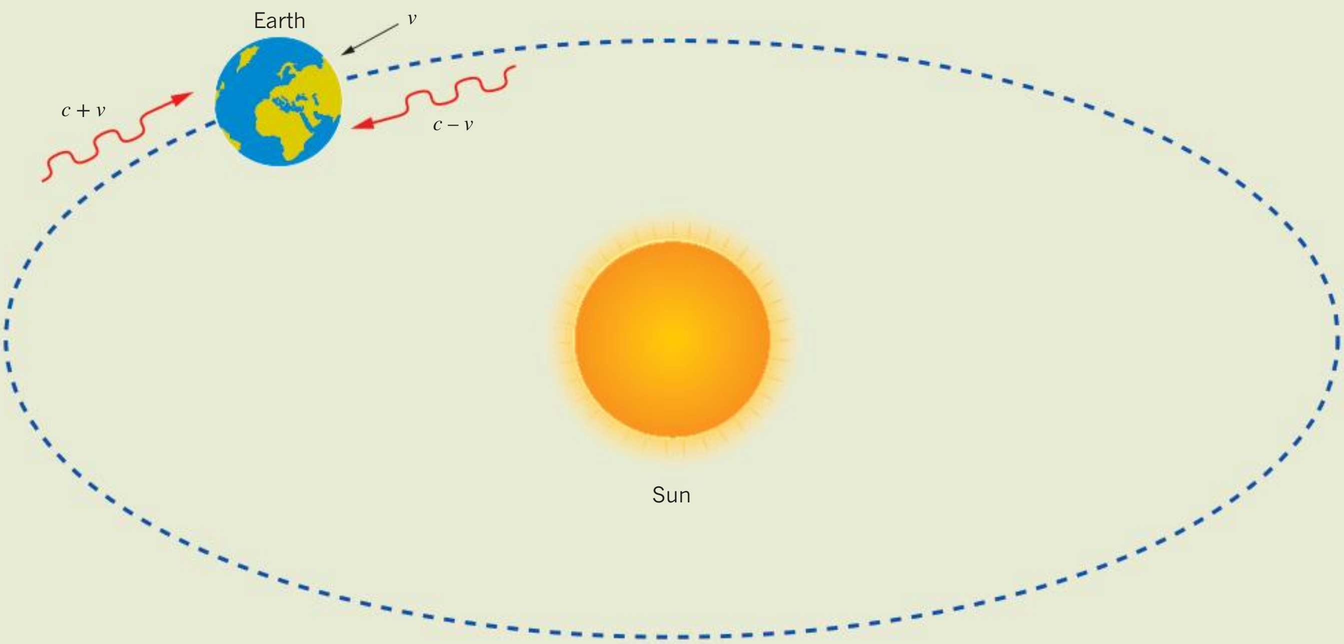


FIGURE 6.1.4 The basic principle of the Michelson–Morley experiment. If the aether is fixed relative to the Sun, and the light is travelling at c relative to the aether and in the same direction as the Earth, the apparent speed should be less than c, i.e. $c - v$. If the light was travelling in the opposite direction to Earth it should appear faster than c, i.e. $c + v$.

Einstein's postulates

The first postulate is basically that of Newton, but Einstein extended it to include the laws of electromagnetism, so elegantly expressed by Maxwell. The second postulate simply takes Maxwell's prediction about the speed of electromagnetic waves in a vacuum at face value.

These two postulates sound simple enough; the only problem was that, according to early Newtonian physics, they were contradictory.

Consider the example illustrated in Figure 6.1.5. Binh is in his spaceship travelling away from Clare at a speed v, and Clare turns on a laser beam to signal Binh. The first postulate seems to imply that the speed of the laser light, as measured by Binh, should be $c - v$, where c is the speed of light in Binh's frame of reference. This is what you would expect if, for example, you were to measure the speed of sound as you travel away from its source; as your velocity gets closer to the speed of sound, the slower the sound waves appear to be travelling.

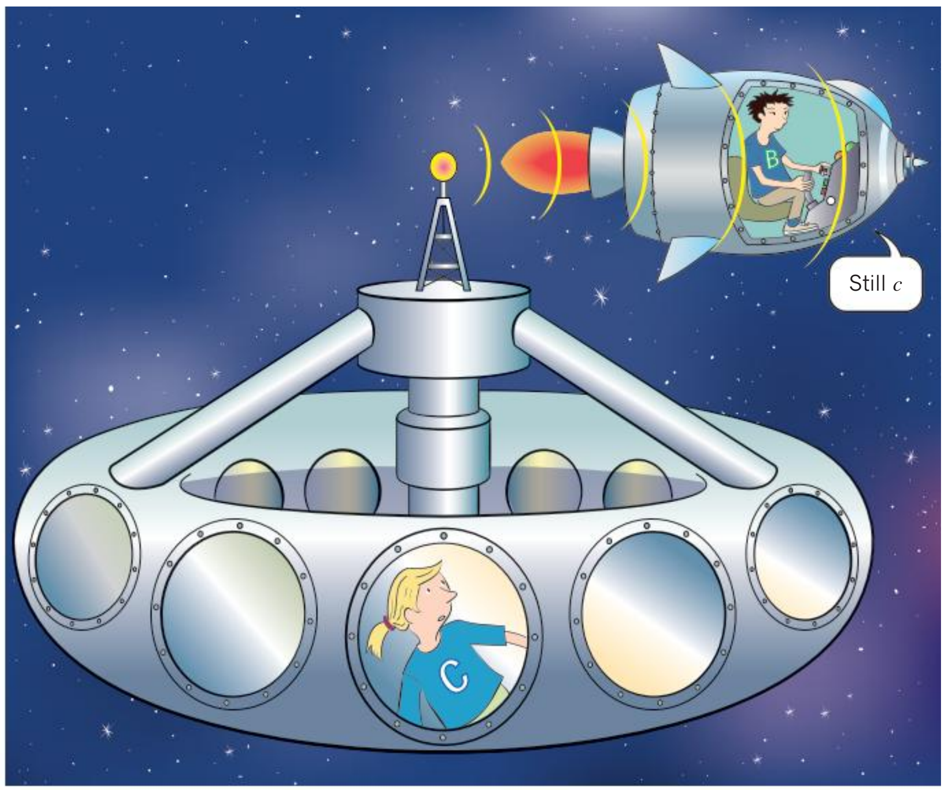


FIGURE 6.1.5 Einstein's two postulates are seemingly contradictory. His first postulate indicates that the speed of the laser light, as measured by Binh, should be $c - v$, whereas his second postulate indicates it should be c. Einstein revisited Newton's assumptions to resolve this problem.

The second postulate, however, tells you that when Binh measures the speed of Clare's laser light, he will find it to be c; that is, 3.00×10^8 m s^{-1}. So at first glance, these two postulates appear to be mutually exclusive. To resolve this problem, Einstein went back to the assumptions on which Newton based his theories.

Newton's assumptions

In 1687, Isaac Newton published his famous *Principia*. At the start of this incredible work, which was the basis for all physics in the following two centuries and beyond, he notes the following assumptions.

The following two statements are assumed to be evident and true:

- *Absolute, true, and mathematical time, of itself, and from its own nature, flows equably without relation to anything external.*
- *Absolute space, in its own nature, without relation to anything external, remains always similar and immovable.*

Newton based all of his laws on these two assumptions: that space and time are constant, uniform and straight. So according to Newton, space is like a big set of *xyz* axes that always have the same scale, and in which distances can be calculated exactly according to Pythagoras' theorem. You expect a metre rule to be the same length whether it is held vertically or horizontally, north–south or east–west, in your classroom or in the International Space Station.

In this space, time flows on at a constant rate, which is the same everywhere. One second in Perth is the same as one second in Melbourne, and one second on the ground is the same as one second up in the air.

Einstein realised that the assumptions that Newton made may not be valid, at least not on scales involving huge distances and speeds approaching the speed of light. The only way in which Einstein's two postulates can both be true is if both space and time are not fixed and unchangeable.

Einstein's *Gedanken* train

To illustrate the consequences of accepting the two postulates he put forward, Einstein discussed a simple thought experiment. It involves a train, moving at a constant velocity.

Amaya and Binh have boarded Einstein's *Gedanken* train and Clare is outside on the platform (refer to Figure 6.1.6). This train has a flashing light bulb set right in the centre of the carriage. Amaya and Binh are watching the flashes of light as they reach the front and back walls of the carriage. They find that the flashes reach the front and back walls at the same time, which is not surprising. Outside, Clare is watching the same flashes of light. Einstein was interested in when Clare saw the flashes reach the end walls.

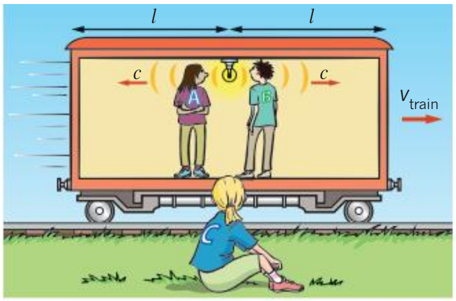


FIGURE 6.1.6 Amaya and Binh see the light take the same time, $\frac{l}{c}$ seconds, to reach the front and back walls.

To appreciate Einstein's ideas, you need to contrast them with what you would normally expect. Consider a situation in which Amaya and Binh are rolling balls towards opposite ends of a train carriage. It is important to appreciate that, while Clare, the outside observer, sees the ball's velocity differently from Amaya and Binh, the times at which various events (balls hitting the ends of the carriage) occur must be the same.

If you had discussed a pulse of sound waves travelling from the centre of the train, you would find exactly the same result: Clare always agrees with Amaya and Binh that the time taken for balls, or sound waves, to reach the end walls is the same. But what about light?

Einstein's second postulate tells you that all observers see light travel at the same speed. Amaya, Binh and Clare will all see the light travelling at 3.00×10^8 m s^{-1}; they do not add or subtract the speed of the train.

If Clare sees the light travelling at the same speed in the forward and backward directions, she will see the light reach the back wall first (refer to Figure 6.1.7). This is because that wall is moving towards the light, whereas the front wall is moving away from the light, and so the light will take longer to catch up to it. This is against the principles of Newtonian physics. Amaya and Binh saw the light flashes reach the ends of the carriage at the same time; Clare saw them reach the walls at different times.

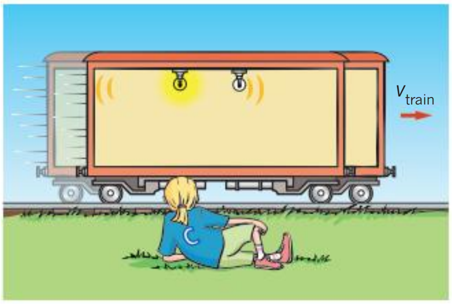


FIGURE 6.1.7 Clare sees the light reach the back wall first, and then the front wall.

The idea that two events that are **simultaneous** (occur at the same time) for one set of observers but are not simultaneous for another is outrageous.

Simultaneity and spacetime

The big difference between the situation for light, and that for balls or sound, is the strange notion that both sets of observers see the speed of light as exactly the same. The velocity of a thrown ball or the velocity of sound in Amaya and Binh's frame of reference will always be different from that in Clare's frame of reference by exactly the velocity of the train. For light, however, there is no difference. As a result, events that are simultaneous for one set of observers are *not simultaneous* for the others. This is a very strange situation that is referred to as a lack of simultaneity.

PHYSICSFILE

Measurement in a thought experiment

The people in the *Gedanken* train would need extremely good measuring devices, such as an atomic clock, and amazingly quick reflexes to take their measurements.

Under normal circumstances, there is no chance of detecting the lack of simultaneity of light beams hitting the front and back walls of a train. This is because the differences in time are about a millionth of a microsecond, well beyond the capacity of even the best stopwatches. The reflexes required to see the light reach the back wall, then see the light encounter the front wall, would also be beyond human ability.

FIGURE 6.1.8 The famous clock tower in Bern, Switzerland, near Einstein's apartment. Its hands move at one minute per minute, but only in the same frame of reference as the clock.

While Einstein's *Gedanken* experiments are purely hypothetical, other experiments based on these ideas are well within the capacity of modern experimental physics. In all cases they confirm Einstein's ideas to a high degree of accuracy.

Einstein said that the only reasonable explanation for how two events that were simultaneous to one set of observers were not simultaneous to another, is that time itself is behaving strangely. The amount of time that has elapsed in one frame of reference is not the same as that which has elapsed in another (see Figure 6.1.8).

In the example shown in Figure 6.1.7, Amaya and Binh saw the light flashes that went forward and backward take the same time to reach the walls. In Clare's frame of reference the times were different. Time, which has one dimension, seems to depend on the frame of reference in which it is measured, and a frame of reference is just a way of defining three-dimensional space. Clearly time and space are somehow interrelated. This four-dimensional relationship, which includes the three dimensions of space and the one dimension of time, is called **spacetime**. Special relativity is all about spacetime.

This was a profound shock to the physicists of Einstein's time. Many of them refused to believe that time was not the constant and unchanging quantity that it was assumed always to have been. And to think that it might 'flow' at a different rate in a moving frame of reference was too mind-boggling for words. That could mean that if you went for a train trip, your clock would go slower, and you would come back having aged slightly less than those who stayed behind.

Einstein's idea was that time and distance are relative. They can have different values when measured by different observers. Simultaneous events in one frame of reference are not necessarily simultaneous when observed from another frame of reference. This is difficult to comprehend at first and will take some time to fully appreciate. Our basic understanding of time and distance (and perhaps mass too) need adjustment when objects travel close to the speed of light. A certain observer might see light travelling through a distance d in a time t at a speed c. A different observer might see light travelling through a different distance, d', in a different time, t', but still at the same speed, c.

Probably because of the tiny differences involved and the highly abstract nature of the work, many physicists simply disregarded the concepts and got on with their work. They thought it could never have any practical results.

OBSERVATIONS THAT NEWTON'S LAWS CAN'T EXPLAIN

With the invention of more accurate measuring devices for time and distance, it became evident that some of the measurements of events didn't agree with the predicted values. These predicted values were based on Newton's laws acting in a framework of Galilean relativity.

Atomic clocks

Measuring time is an exercise in precision, replicating an interval of one second over and over again, until 86 400 of them equals the time for one rotation of the Earth, or one day. There have been many mechanical solutions to this problem in the past using cogs and levers, weights and dials. The accuracies of these devices varied, with some of them gaining or losing seconds or minutes per day.

To correct your clock you would need to frequently adjust it against a standard clock. To help in this recalibration, radio stations would broadcast a time signal, so you could set your clock each day. Typically they would broadcast a series of five beeps counting down to each hour. You could also phone a number that would tell you 'at the tone it will be six o'clock … beep'.

For scientists, clocks with this level of inaccuracy could only be reliable for measuring events to one or two decimal places, which is fine for verifying relatively slow motion. Such clocks could not differentiate between two events occurring over a much smaller time interval.

Before 1967, the standard of one second was based on a fraction of the time it took for the Earth to orbit the Sun, a far-from-ideal standard. From 1967 onwards, the basis for the unit of time was changed to be a certain number of transitions of the outermost electron of a caesium-133 isotope. In fact, one second is now defined as: 9 192 631 770 oscillations of the 6s electron of the Cs-133 isotope. The remarkable precision of this oscillation resulted in atomic clocks (the first of which is shown in Figure 6.1.9) with an accuracy of 1 second in 1.4 million years, and the ability to measure time to an incredible number of decimal places. It is at these levels of measurement that the predictions of Newton's laws of motion vary from the measured values.

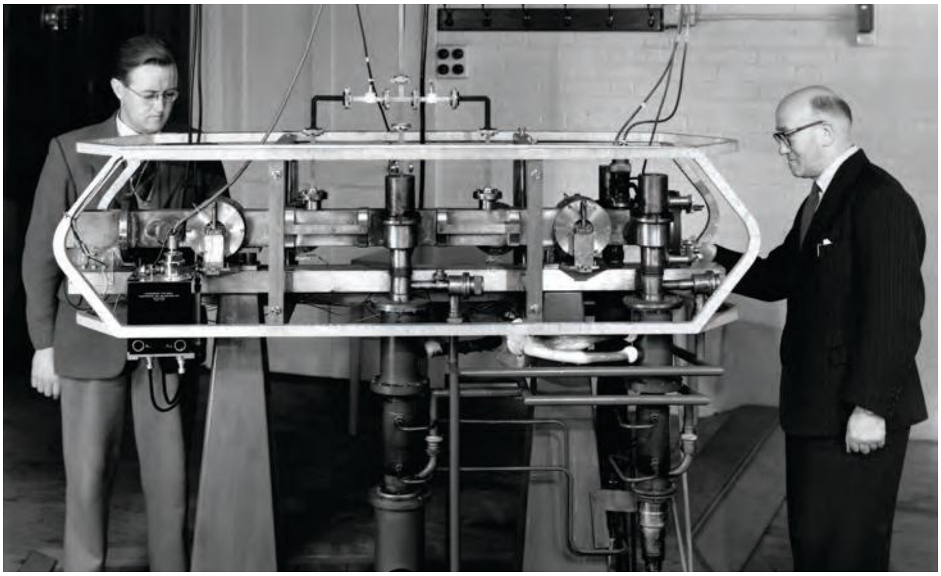


FIGURE 6.1.9 The first atomic clock, developed in 1955, and used to set the standard of one second from 1967.

LONG-LIVED MUONS

When certain unstable particles (like pions, which have a precisely known decay rate) are accelerated to almost the speed of light, their life spans are measured to be longer than when the particles are stationary. For example, the mean lifetime of the positive pion, π^+, is 0.000 000 026 033 s (26.033 ns) when it is stationary relative to the atomic clock that is measuring it. However, when it is moving at 99% of the speed of light, its mean lifetime as measured by the stationary atomic clock is 184.54 ns. This means that the moving pion exists seven times longer than a stationary pion.

In the Earth's atmosphere, high-energy cosmic rays interact with the nuclei of oxygen atoms 15 km above the surface of the Earth to create a cascade of high-velocity sub-atomic particles. One of these particles is a muon, which is unstable. The mean lifetime of a stationary muon, as measured by an atomic clock, is 0.000 002 196 s (2.196 μs). The muons created by cosmic radiation typically travel at 99.97% of the speed of light, so at this speed Newtonian physics would predict that a muon would travel about 659 m:

$$\begin{aligned} s &= v\Delta t \\ &= 0.9997 \times 3 \times 10^{8} \times 2.196 \times 10^{-6} \\ &= 658.6 \text{ m (or roughly 659 m)} \end{aligned}$$

After 10 lifetimes, you can expect there to be essentially no muons remaining. So after beginning at a height of 15 km and travelling through a distance of 6.58 km, to a height of about 8.42 km above the surface of the Earth, you would expect that no muons would be detected.

However, muons created by cosmic radiation are actually detected at the surface of the Earth. This means that the fast-moving muons have existed for a much longer period of time than they should have. A muon that strikes the surface would have existed at least 22.8 times its predicted lifespan as a stationary muon, based on Newtonian physics. Once again, Newtonian physics and Galilean relativity cannot explain this observation. The next section, Time dilation, will explain why this happens.

PHYSICS IN ACTION

Particles gaining mass

When an object moves in a circular path, it does so as a result of a centripetal force that acts towards the centre of the circular path. Centripetal force therefore acts continuously at a right angle to the velocity of the object. There are a number of actions that could cause the centripetal force on an object, such as the tension in a string tied to a rubber stopper or the gravitational force of the Earth on the Moon.

Another action that causes circular motion is the force on a charged particle that is moving at right angles to a magnetic field. The equation that represents the relationship between the magnetic force (F_B) and the centripetal force (F_C) is:

$$F_B = F_C$$

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

$$r = \frac{mv}{qB}$$

$$r = \frac{m}{qB} \times v$$

The final relationship shows that, if the mass m, charge q and magnetic field B are all constant, then the radius of the circular path is directly proportional to the velocity of the charged particle. So, theoretically, if the velocity increases by a factor of 2, then the radius will also increase by a factor of 2. However, this is not the case.

In circular accelerating devices, such as cyclotrons and the Australian Synchrotron, it is evident that, as the velocity of a charged particle increases, the radius of its path also increases, but to a much greater degree than that expected. According to the relationship shown above, if the charge q and the magnetic field B don't change, then the only explanation for the extra increase in radius is that the mass of the particle m must increase.

In fact, the mass of an electron travelling at 99.999 99% of the speed of light seems to increase to 6000 times the mass of an electron at rest. There is no explanation for this phenomenon in Galileo's relativity or Newtonian physics. This will all be explained later in the chapter, using Einstein's theories.

6.1 Review

SUMMARY

- Einstein decided that Galileo's principle of relativity was so elegant it simply had to be true, and he was also convinced that Maxwell's electromagnetic equations, and their predictions, were sound.
- Einstein's two postulates of special relativity can be abbreviated to:
  - I The laws of physics are the same in all inertial frames of reference.
  - II The speed of light is the same to all observers.
- Einstein realised that accepting both of these postulates implied that space and time were not absolute and independent, but were related in some way.
- Two events that are simultaneous in one frame of reference are not necessarily simultaneous in another.
- This implies that time measured in different frames of reference might not be the same. Time and space are related in a four-dimensional universe of spacetime.
- Observations of the lifetimes of sub-atomic particles that are accelerated to high speeds indicate that they exist for longer than when they are stationary.
- High-speed muons created in Earth's upper atmosphere should not last long enough to reach Earth's surface, but they do. The moving muons have longer lifetimes than stationary muons.

KEY QUESTIONS

1 Why did the physicists of the late 19th century feel the need to invent the idea of the aether?
- **A** It was required to satisfy the principle of relativity.
- **B** It was required to satisfy Maxwell's equations.
- **C** They thought that it would be impossible that totally empty space could occur in nature.
- **D** They thought that there should be a medium that carries light waves just as air carries sound waves.

2 Which of the following are reasonably good inertial frames of reference? More than one answer is possible.
- **A** an aircraft in steady flight
- **B** an aircraft taking off
- **C** a car turning a corner
- **D** a car driving up a hill of constant slope at a steady velocity

3 Two spaceships are travelling for a while with a constant relative velocity. Then one begins to accelerate. A passenger with a laser-based velocity measurer sees the relative velocity increase. How could this passenger tell whether it was his own or the other ship that began to accelerate?

4 Tom, who is in the centre of a train carriage moving at constant velocity, rolls a ball towards the front of the train, while at the same time he blows a whistle and shines a laser towards the front of the train. What will Jana, who is on the ground outside the train, observe compared with Tom regarding the speed of the ball, the sound and the light?

5 If the speed of sound in air is 340 m s^{-1}, at what speed would the sound from a fire truck siren appear to be travelling in the following situations?
- **a** You are driving towards the stationary fire truck at 30 m s^{-1}.
- **b** You are driving away from the stationary truck at 40 m s^{-1}.
- **c** You are stationary and the fire truck is heading towards you at 20 m s^{-1}.
- **d** You are driving at 30 m s^{-1} and about to overtake the fire truck, which is travelling at 20 m s^{-1} in the same direction.

6 In order to resolve the apparent conflict resulting from his two postulates, Einstein rejected some of Newton's assumptions. Which of the following statements is a consequence of this?
- **A** Time is not constant in all frames of reference.
- **B** Absolute, true, and mathematical time, of itself, and from its own nature, flows equably without relation to anything external.
- **C** One second in any inertial frame of reference is the same as one second in any other inertial frame of reference.
- **D** Space and time are independent of each other.

7 Anna is at the front end of a train carriage moving at 10 m s^{-1}. She throws a ball back to Ben, who is 5 m away at the other end of the carriage. Ben catches it 0.2 s after it was thrown. Chloe is watching all this from the side of the track.
- **a** At what velocity does Chloe see the thrown ball travelling?
- **b** How far, in Chloe's frame of reference, did the ball move while in flight?
- **c** How long was it in flight in Chloe's frame of reference?

8 Imagine that the speed of light has suddenly slowed down to only 50 m s^{-1} and this time Anna (still at the front of the 5 m train moving at 10 m s^{-1} in Question 7) sends a flash of light towards Ben.
- **a** From Anna's point of view, how long does it take the light flash to reach Ben?
- **b** How fast was the light travelling in Ben's frame of reference?
- **c** In Chloe's frame of reference, how far did the train travel in 0.1 s?
- **d** How fast was the light travelling in Chloe's frame of reference?
- **e** Approximately, when did Chloe see the light reach Ben?

9 Why was the development of atomic clocks important to the advancement of Einstein's special theory of relativity?

10 Complete the following sentences by selecting the correct term in bold.

Muons have **very short/prolonged** lives. On average, muons live for approximately 2.2 **s/μs**. Their speeds are measured as they travel through the atmosphere. A muon's speed is **about a tenth of/very similar to** the speed of light. According to Newtonian laws, muons **should/should not** reach the Earth's surface. However, many **do/do not**.

6.2 Time dilation

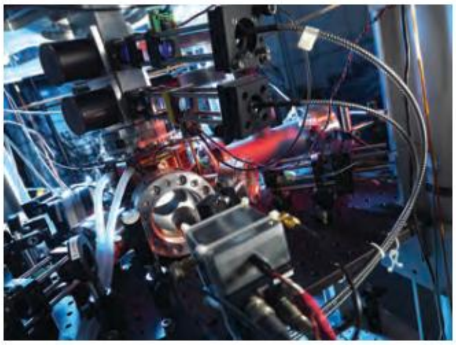


FIGURE 6.2.1 The duration of one second can be measured very precisely using a caesium atomic clock like this one.

Extremely precise atomic clocks like that shown in Figure 6.2.1 enabled very short-lived events to be measured to a large number of decimal places. At this level of precision, some unusual observations were made regarding the life spans of some high-speed sub-atomic particles when compared to the life spans of those same particles at rest. This section explores the concept of time dilation as an explanation for these observations.

TIME IN DIFFERENT FRAMES OF REFERENCE

The consequences of Einstein's two postulates have been discussed, in general terms, when they are applied to a simple *Gedanken* situation, such as a moving train. Observers inside the train see two simultaneous events while those outside see the same two events occurring at different times. Certainly the differences are extremely small and would not be noticeable by an observer in any actual train, unless they had an atomic clock. For aircraft flying at supersonic speeds, the differences, while very small, become measurable by the most precise clocks. For sub-atomic particles, such as pions in accelerators like the Australian Synchrotron, the differences in time become more significant, and so in situations like this, where speeds approach the speed of light, it is important to use calculations that take Einstein's theory into account.

The light clock

If you want to observe time dilation in the moving train or among moving sub-atomic particles, you need to watch a clock in a moving reference frame to see if it is actually going slower. The term 'dilation' in this context means slower.

Consider Amaya and Binh riding in a *Gedanken* spaceship that can travel at speeds close to the speed of light. Clare is going to watch from a space station, which according to Clare is a stationary frame of reference. Amaya and Binh have taken along a clock, which (it is assumed) Clare can read, even from a large distance away.

Like any clock, this clock is governed by a regular oscillation that defines a period of time.

Amaya's *Gedanken* clock has a light pulse that bounces back and forth between two mirrors. One mirror is on the floor and the other on the ceiling, as shown in Figure 6.2.2. When a light pulse oscillates from one mirror to the other and back, you can consider that period of time to be 'one unit'. Clare has an identical clock in her own space station, which she can compare to Amaya's clock.

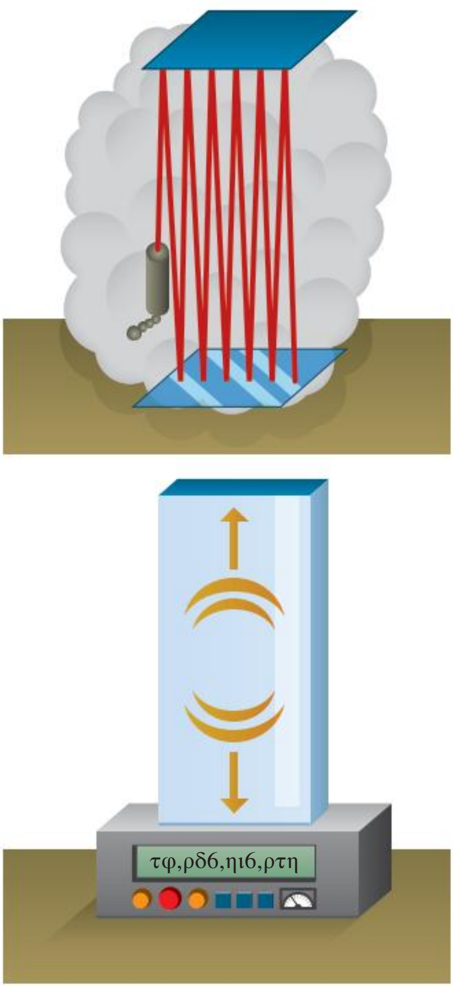


FIGURE 6.2.2 The *Gedanken* light clock 'ticks' each time the light pulse reflects off the bottom mirror.

The advantage of this clock is that it can be used to predict how motion will affect it by using Pythagoras' theorem and some algebra. The clock has been set up in the spaceship so that the light pulses oscillate up and down a distance d that is at right angles to the direction of travel. The distance d is shown by a black arrow in the centre position of the moving spacecraft in Figure 6.2.3. As the spaceship speeds along, the light will trace out a zigzag path, as shown by the red dotted line in Figure 6.2.3.

Only one of the oscillations of the light pulse needs to be considered, as all the other oscillations will have the same geometry. One 'unit of time' will be the time taken for the light pulse to oscillate once. In the frame of reference of the spaceship, Amaya and Binh see a unit of time equal to t_a. Clare, from her frame of reference, will see a different time, t_c. The relationship between these two times will now be determined.

Amaya and Binh see the light pulse travel at the speed of light, c, along the distance $2d$, from the bottom mirror to the top and back again, in time t_a. So the distance that the light pulse travels is given by:

$2d = c \times t_a$

On the other hand, Clare sees the light travel a longer path that is shown as the red dotted line in Figure 6.2.3.

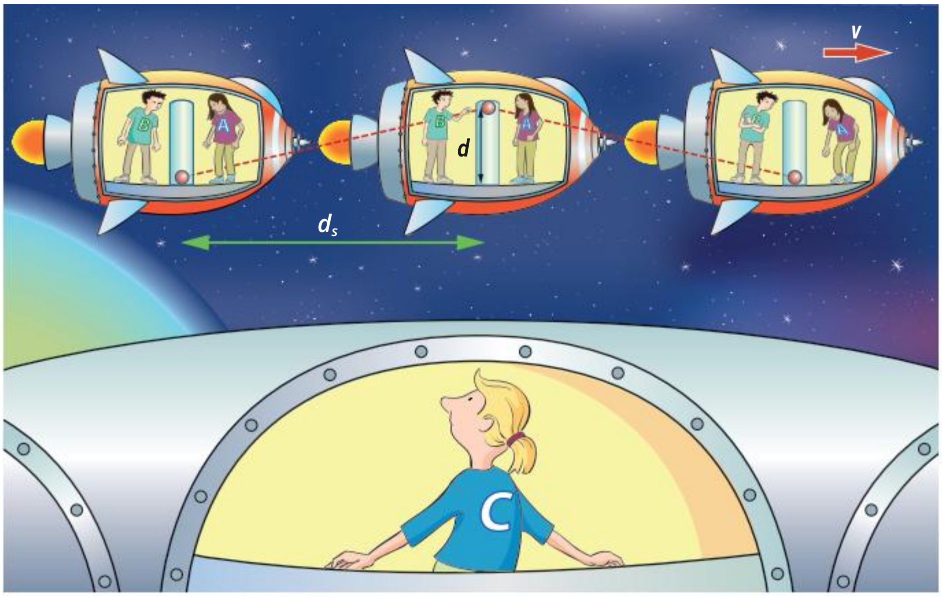


FIGURE 6.2.3 Clare can see that in one unit of time the *Gedanken* light clock 'ticks' each time the light pulse reflects off the bottom mirror. She also sees that the light pulses travel a zigzag path between the mirrors.

The ship moves with a speed v, and so in one unit of time as seen by Clare, t_c, the spaceship will travel a distance $2 \times d_s$, equal to the velocity multiplied by the time taken for her to see one oscillation:

$$2d_s = v \times t_c$$

Consider only half of the light oscillation for now. The light pulse not only travels the vertical distance d in the clock, but also travels forward as the spaceship moves through the distance d_s, making the combined distance d_c. Therefore, according to Pythagoras's theorem:

$$d_c^2 = d^2 + d_s^2$$

$$d_c^2 = d^2 + \left(\frac{vt_c}{2}\right)^2$$

$$d_c = \sqrt{\left(d^2 + \left(\frac{vt_c}{2}\right)^2\right)}$$

Clare sees this light pulse travelling twice this combined distance at the speed of light c, in a period of time t_c measured on her clock. So:

$$2d_c = c \times t_c$$

Equating and rearranging the two expressions for d_c gives:

$$\frac{c \times t_c}{2} = \sqrt{\left(d^2 + \left(\frac{vt_c}{2}\right)^2\right)}$$

$$c \times t_c = 2 \times \sqrt{\left(d^2 + \left(\frac{vt_c}{2}\right)^2\right)}$$

$$c \times t_c = \sqrt{\left(4d^2 + 4 \times \left(\frac{vt_c}{2}\right)^2\right)}$$

$$t_c = \frac{\sqrt{4d^2 + (vt_c)^2}}{c}$$

From Amaya and Binh's frame of reference, where they see the light pulse travelling a distance $2d$ at speed c in a time t_a, the previously given equation can be rewritten in terms of d as:

$$d = \frac{c \times t_a}{2}$$

Note that you have used the same value for c in both of these equations, something you would never do in **classical physics**, but something Einstein insists you must.

Substituting this expression for d into the previous equation gives:

$$t_c = \frac{\sqrt{4 \times \left(\frac{ct_a}{2}\right)^2 + (vt_c)^2}}{c}$$

Now square both sides and simplify to make t_c^2 the subject:

$$t_c^2 = \frac{\frac{4(ct_a)^2}{4} + (vt_c)^2}{c^2}$$

$$t_c^2 = \frac{c^2 t_a^2 + v^2 t_c^2}{c^2}$$

$$t_c^2 = \frac{c^2 t_a^2}{c^2} + \frac{v^2 t_c^2}{c^2}$$

$$t_c^2 = t_a^2 + \frac{v^2 t_c^2}{c^2}$$

Group the terms with t_c^2 together and factorise:

$$t_c^2 - \frac{v^2 t_c^2}{c^2} = t_a^2$$

$$t_c^2 \left(1 - \frac{v^2}{c^2}\right) = t_a^2$$

Take the square root of both sides and make t_c the subject:

$$t_c \sqrt{\left(1 - \frac{v^2}{c^2}\right)} = t_a$$

$$t_c = \frac{t_a}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$$

As v can never be larger than c, the denominator in the equation above must be less than one. Any number divided by a number less than one must result in a larger number, so $t_c > t_a$

This final equation shows that the time that Clare measures, t_c, is greater than the time that Amaya and Binh measure, t_a, for the same event.

PHYSICSFILE

The zigzag path of light

Mathematically, you can see that time dilation results from the strange behaviour of light. As light travels on the diagonal zigzag path, it does so at speed c, not at a faster speed resulting from the additional component of the spaceship's motion as, for example, would be true for a boat zigzagging across a river as it is carried along by the current.

TIME DILATION

In Einstein's equation for time dilation, the symbol t is used to represent the time that a stationary observer (Clare) measures using a stationary clock, for an event that the observer sees occurring in a moving frame of reference. The symbol t_0 is then the time that passes on the moving clock, which is also known as the **proper time**.

The factor that the proper time is multiplied by is given the symbol gamma, γ, so that:

$$t = t_0 \gamma$$

where $\gamma = \frac{1}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$

v is the speed of the moving frame of reference

c is the speed of light in a vacuum (3×10^8 m s^{-1})

t is the time observed in the stationary frame

t_0 is the time observed in the moving frame (proper time)

The physicist H. A. Lorentz first introduced the factor γ in an attempt to explain the results of the Michelson–Morley experiment, so it is often known as the **Lorentz factor**.

Table 6.2.1 and Figure 6.2.4 show the effect of varying the value of v on the value for γ.

v (m s^{-1})	$\frac{v}{c}$	γ
3.00×10^2	0.000001	1.000000000
3.00×10^5	0.00100	1.0000005
3.00×10^7	0.100	1.005
1.50×10^8	0.500	1.155
2.60×10^8	0.866	2.00
2.70×10^8	0.900	2.29
2.97×10^8	0.990	7.09
2.997×10^8	0.999	22.4

TABLE 6.2.1 The value of the Lorentz factor at various speeds.

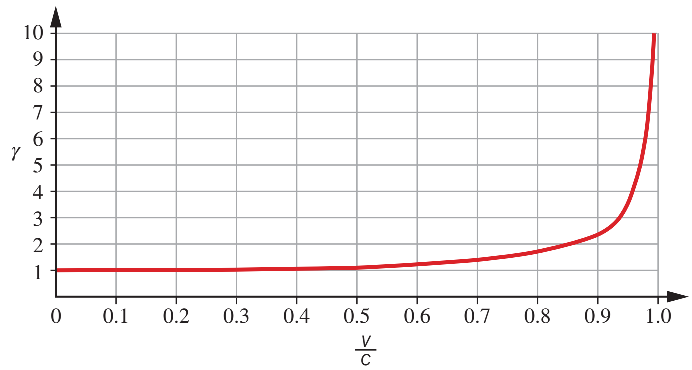


FIGURE 6.2.4 The graph of the Lorentz factor versus $\frac{v}{c}$.

EXTENSION

The Lorentz factor

The Lorentz factor becomes so close to 1 for values of v less than about $0.0001c$ that the expression can't be used with a normal calculator. Fortunately, there is a simple way to find the value of γ for speeds less than about 1% of c. The binomial expansion of the term $(1 - x)^n$ tells you that:

$(1 - x)^n \approx (1 - nx)$ provided that: $x << 1$

For the Lorentz factor this means that:

$x = \left(\frac{v}{c}\right)^2$ and $n = -\frac{1}{2}$

and so:

$$\gamma = 1 + \frac{1}{2}\left(\frac{v}{c}\right)^2$$

Thus the part of the factor that is greater than 1 can simply be found from the term:

$$\frac{1}{2}\left(\frac{v}{c}\right)^2$$

Sometimes it is useful to make v the subject in the equation for the Lorentz factor. This produces:

$$v = c\sqrt{\left(1 - \frac{1}{\gamma^2}\right)}$$

From the data in Table 6.2.1, a velocity of 300 m s^{-1} results in a Lorentz factor of essentially 1. So for relatively low-speed spaceships, a stationary observer measures the oscillation of light in the light clock on the spaceship to be the same as in her own stationary light clock. This implies that time is passing at essentially the same rate in both frames of reference.

When the spaceship is travelling at $0.990c$, a stationary observer like Clare will measure that a single oscillation of light in the spaceship's light clock will take seven oscillations of her own stationary light clock. According to Clare, time for the objects and people in the moving frame of reference has slowed down to one-seventh of 'normal' time.

As the speed approaches the speed of light, time in the moving frame, as viewed from the stationary frame, slows down more and more. So, if you were able to see the clock travelling on a light wave, the clock would not be 'ticking' at all. In other words, time would be seen to stand still.

It is important to realise that Amaya and Binh do not perceive their time slowing down at all. To them, their clock keeps ticking away at the usual rate and events in their frame of reference take the same time as they normally would. It is the series of events that Clare sees and measures in Amaya and Binh's frame that go slowly. Binh and Amaya are moving in slow motion because, according to Clare's observations, time for them has slowed down (see Figure 6.2.5).

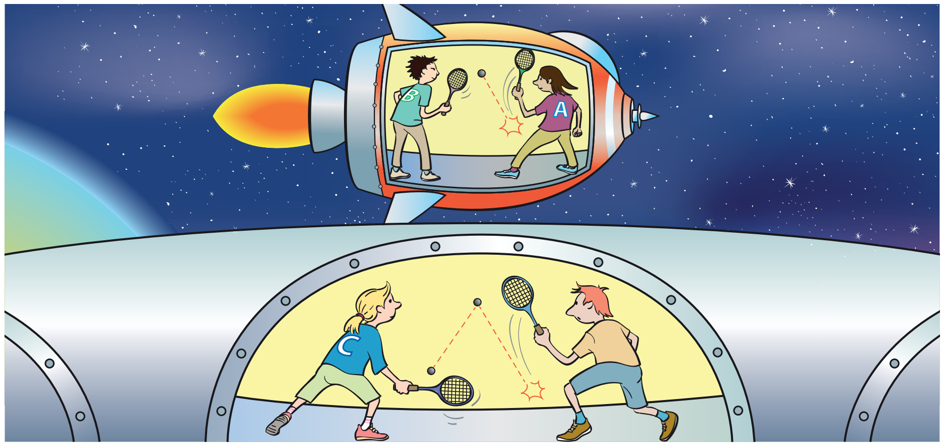


FIGURE 6.2.5 As Clare watches Amaya and Binh play space squash, the ball seems to be moving much more slowly than in her own game.

Worked example 6.2.1

TIME DILATION

Assume *Gedanken* conditions exist in this example. A stationary observer on Earth sees a very fast car passing by, travelling at 2.50×10^8 m s^{-1}. In the car is a clock on which the stationary observer sees 3.00 s pass. Calculate how many seconds pass by on the stationary observer's clock during this observation. Use $c = 3.00 \times 10^8$ m s^{-1}.

Thinking	Working
Identify the variables: the time for the stationary observer is t, the proper time for the moving clock is t_0 and the velocities are v and the constant c.	$t = ?$ $t_0 = 3.00$ s $v = 2.50 \times 10^8$ m s^{-1} $c = 3.00 \times 10^8$ m s^{-1}
Use Einstein's time dilation formula and the Lorentz factor.	$t = t_0\gamma$ $= \frac{t_0}{\sqrt{1 - \frac{v^2}{c^2}}}$
Substitute the values for t_0, v and c into the equation and calculate the answer, t.	$t = \frac{3.00}{\sqrt{1 - \frac{(2.50 \times 10^8)^2}{(3.00 \times 10^8)^2}}}$ $= \frac{3.00}{0.55277}$ $= 5.43$ s

Worked example: Try yourself 6.2.1

TIME DILATION

Assume *Gedanken* conditions exist in this example. A stationary observer on Earth sees a very fast scooter passing by, travelling at 2.98×10^8 m s^{-1}. On the wrist of the rider is a watch on which the stationary observer sees 60.0 s pass. Calculate how many seconds pass by on the stationary observer's clock during this observation. Use $c = 3.00 \times 10^8$ m s^{-1}.

LOOKING BACK TO THE STATIONARY OBSERVER

So far you have been looking at the situation from Clare's point of view, not Amaya's and Binh's. Galileo had said that all inertial frames of reference are equivalent. It follows then that, according to Amaya and Binh, as they look out their window at Clare in her space station receding from them, they can consider that it is they who are at rest and it is Clare and her space station that are moving away at a velocity near the speed of light. This is what Galileo's principle of relativity and Einstein's first postulate are all about.

If Amaya and Binh watch the light clock in Clare's space station, they see that time has slowed down for Clare, as they would observe Clare's moving light clock oscillation taking longer than their stationary light clock oscillation. This raises the question: Whose time actually runs slowly?

The answer is that they are both right. The whole point of relativity is that you can only measure quantities relative to some particular frame of reference, not in any absolute sense. Certainly Amaya and Binh see Clare as though in slow motion and Clare sees them in slow motion. Remember that there is no absolute frame of reference and so there is no absolute clock ticking away the absolute 'right' time. All that you can be sure of is that time in your own inertial frame of reference is ticking away at a rate of one second per second.

THE TWIN PARADOX

If Clare sees time for Amaya and Binh running slowly, then Amaya and Binh will age slowly. But if Amaya and Binh see that time for Clare has slowed down, then Clare will age more slowly. So what happens when Amaya and Binh decide to turn their spaceship around and come home? Who will have aged more?

To solve this **paradox**, or contradiction, Einstein described a thought experiment in which one of a set of twins heads off on a long space journey, while the other twin stays on Earth.

The travelling twin finds that when she returns, her remaining twin has become quite elderly (refer to Figure 6.2.6). While each twin is in constant motion relative to the other, they both see the other twin ageing more slowly. So why did the twin on the spaceship age more slowly than the twin on Earth?

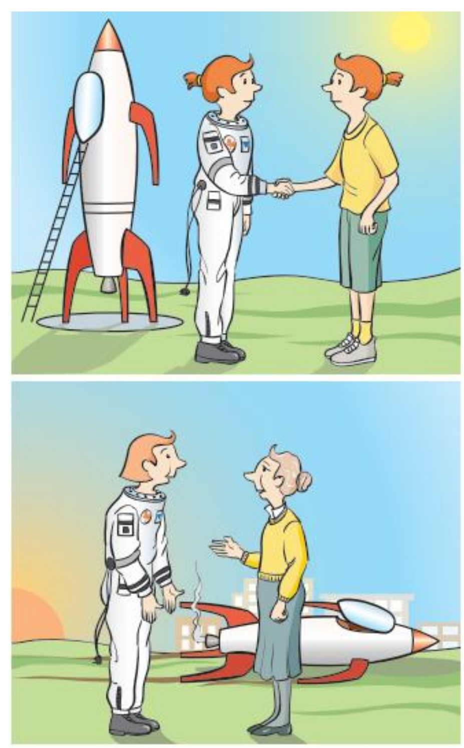


FIGURE 6.2.6 The Twin Paradox describes the phenomenon where one twin ages less quickly than the other after travelling in a non-inertial frame.

The key to this apparent paradox is that only one twin has spent the entire time in an inertial (non-accelerating) frame of reference. The other twin spent some time in non-inertial frames of reference. The twin that got on the spaceship accelerated away from Earth, decelerated as she slowed down, then accelerated back towards Earth, and finally decelerated as she slowed down to land back on Earth. It is the acceleration that makes all the difference.

It is important to point out that Einstein's 1905 theory of relativity deals only with frames of reference that are in constant motion, that is, inertial frames of reference. For this reason, it is called the theory of *special* relativity. Special relativity does not deal with accelerated frames of reference. Ten years later, Einstein put forward the theory of *general* relativity, which does deal with situations in which acceleration occurs; that is, non-inertial frames of reference. As part of this theory, he showed that in an accelerated frame of reference, time also slows down.

If you apply the twin paradox situation to Amaya, Binh and Clare, as Clare watched from her inertial frame of reference, the general theory of relativity tells you that her view of Amaya and Binh in the non-inertial frame shows them ageing slowly. During this time, Amaya and Binh see Clare's time passing quickly. As a result, they will see Clare age more rapidly while they are accelerating, and more slowly when they are travelling at constant velocity. Clare sees Amaya and Binh aging slower and slower as they gain speed, then aging constantly but slowly as they travel at a constant speed. Amaya and Binh never age rapidly.

But how do you know that it is Amaya and Binh that have accelerated and not Clare, because that is what it would look like for Amaya and Binh looking out of their window at Clare? For the answer to this you need to ask Amaya and Binh if they noticed anything unusual in their frame of reference, for example did the surface of the water in their bottles tilt at an angle to the horizontal, or did the handles hanging down from the ceiling lean forwards or backwards. If you asked Clare these questions she would say no, while Amaya and Binh would say yes. So it was Amaya and Binh that accelerated and not Clare.

Although it is often called a paradox, there is actually nothing impossible or illogical about this story. Einstein himself pointed out that, due to the Earth's rotation, and therefore centripetal acceleration, a clock on the Earth's equator would run a little more slowly than one at the poles. This has now actually been found to be the case. In fact, in 1971 accurate atomic clocks were flown around the world on commercial flights. When compared with those left behind, the difference of about a quarter of a microsecond was just what Einstein's theory predicted. Now there are many satellites in orbit around the Earth, so the theory has been well and truly tested many times. Indeed, global positioning systems (GPS) must take the relativistic corrections into account to ensure their accuracy.

PHYSICSFILE

Is light slowing down?

Recently there has been publicity given to research that has suggested that the speed of light is slowing down. Some have even suggested that Einstein's theory of relativity itself is under threat. The research, based on analysis of light from very distant quasars, actually suggests that there have been very small changes in what is called the fine structure constant, which is made up of three more basic constants: the speed of light, the charge on an electron and Planck's constant.

Prominent theoretical physicist Professor Paul Davies and others have suggested that if the evidence is correct, then it is probably the speed of light that is changing. If proved correct, no doubt this new data will modify some aspects of relativity, but to suggest that it will overturn relativity is a wild exaggeration.

EXPLAINING HIGH-ALTITUDE MUONS

In Section 6.1 'Einstein's theory of special relativity', the surprising observation of high-speed muons originating 15 km up in the atmosphere and yet reaching the surface of the Earth was discussed. It could only be explained if the mean lifetime of the short-lived particles were extended far beyond their normal mean lifetime.

Time dilation provides the explanation to this unusual observation.

The 'normal' mean lifetime of a muon is about 2.2 μs. However, this is the mean lifetime when measured in a stationary frame of reference. Muons travel very fast; in fact a speed as great as $0.999c$ is very possible. At this speed, an observer on Earth will see the lifetime of a muon as far greater than 2.2 μs:

$$t = t_0\gamma$$

$$= \frac{t_a}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$= \frac{2.2 \times 10^{-6}}{\sqrt{1 - \frac{(0.999c)^2}{c^2}}}$$

$$= \frac{2.2 \times 10^{-6}}{\sqrt{1 - 0.999^2}}$$

$= 49.21$ μs (which is more than 22 times as long as in the stationary frame!)

An observer on Earth would see the muon's time run much slower. The slower time means that many muons live long enough to reach the Earth's surface.

6.2 Review

SUMMARY

- The pulses in a light clock in a moving frame of reference have to travel further when observed from a stationary frame.
- Because of the constancy of the speed of light, this effectively means that time appears to have slowed in a moving frame.
- Time in a moving frame seems to flow more slowly according to the equation: $t = t_0\gamma$
- where t_0 is the time in the moving frame (proper time), t is the time observed from the stationary frame and γ is the Lorentz factor:

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

- Observers in relative motion both see time slowing in the other frame of reference; that is, each sees the other ageing more slowly.
- If one observer accelerates in order to return to meet the other, then that accelerated observer will have aged less than the other.
- Time dilation provides the explanation for the occurrence of muons reaching the Earth's surface after originating 15 km up in the upper atmosphere, when they should all decay within 7 km of their journey according to classical physics.

KEY QUESTIONS

For the following questions, assume Gedanken conditions exist and let $c = 3 \times 10^8$ m s^{-1} unless stated otherwise.

1 Complete the following sentences by selecting the correct term from those in bold.
In a device called a **light/mechanical/digital** clock, the **speed/oscillation/wavelength** of light is used as a means of measuring **time/mass**, as the speed of light is **unknown/variable/constant** no matter from which inertial frame of reference it is viewed.

2 To what does the term 'proper time', t_0, refer?

3 An observer is standing on a train platform as a very fast train passes by at a speed of 1.75×10^8 m s^{-1}. The observer notices the time on a passenger's phone as the passenger drops the phone to the floor. According to the clock on the phone, it takes 1.05 s to hit the floor. Calculate how much time has passed on the platform's clock during this time.

4 An observer standing on a comet is watching as a satellite approaches at a speed of 2.30×10^8 m s^{-1}. The observer times on her watch that the solar panels on the satellite unfold in 75.0 s. Calculate how much time the observer sees as having passed on the satellite's clock.

5 A student standing by the side of a road sees a very fast MG sports car driving past. The driver times on his car's clock that it takes 5.50 s for the student to pick up her bag. If the MG is moving at a speed of 2.75×10^8 m s^{-1}, calculate how much time the driver sees has passed on the student's watch as she picks up the bag.

6 If Anna saw Ben fly by at 0.5*c*, how long, in her frame, would it take Ben's clock to tick 1 second?

7 Anna's *Gedanken* light clock has a height of 1 m between the mirrors, and relative to Chloe her spaceship is travelling at 90% of the speed of light ($c = 3.0 \times 10^8$ m s^{-1}). One tick is the time for light to go from one mirror to the other.

- **a** How far does the light flash travel in Anna's frame of reference in one tick, t_A?
- **b** What is the tick time, t_A, for the clock in Anna's frame?

As the light takes a zigzag path in her frame, Chloe sees the clock ticking at a slower rate, t_c.

- **c** In terms of *c* and t_c what is the length of the zigzag path that the flash travels in one tick in Chloe's frame?
- **d** What is the tick time of the clock in Chloe's frame?
- **e** What is the ratio of Chloe's tick to Anna's tick?

8 A muon created at an altitude of 15.0 km above the Earth is moving at a speed of 0.992 times the speed of light. The mean lifetime of a muon at rest is 2.20×10^{-6} s.

- **a** Calculate the lifetime of the moving muon as timed by a stationary observer.
- **b** Using classical physics equations and the results from part a, calculate the non-relativistic distance and the relativistic distance travelled by the moving muon during one lifetime.

9 A high-speed, sub-atomic particle is accelerated by a linear accelerator to a speed of 2.83×10^8 m s^{-1}. A researcher measures that it only leaves, on average, a track that is 2.50 cm long in the bubble chamber. Calculate the mean lifetime of the same particle if it were at rest relative to the researcher and her timer.

10 Briefly explain why Einstein said that a clock at Earth's equator should run slightly slower than one at the Earth's poles. Why do we not find this to be a problem?

6.3 Length contraction

The previous section described how time can only be measured relative to some particular frame of reference, but not in any absolute sense. Because of the constancy of the speed of light, this effectively means that time appears to have slowed in a moving frame relative to the frame of an observer. Einstein describes how space and time are interrelated, so it follows then that space, and therefore length, is not absolute (Figure 6.3.1). This section explores the effect on the length of an object based on the motion of the object in an inertial frame of reference.

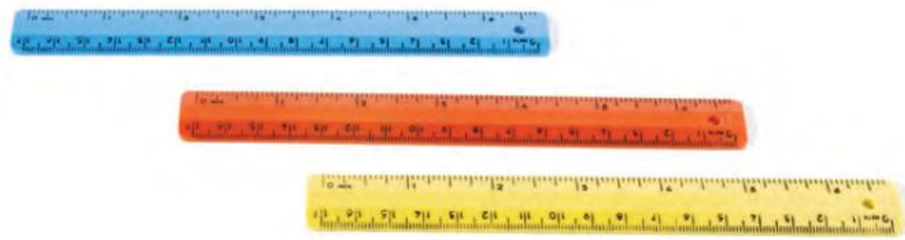


FIGURE 6.3.1 Length is relative to the frame of reference and the direction of motion.

LENGTH IN DIFFERENT INERTIAL FRAMES

You already have a clue to the fact that lengths depend on who is doing the measuring and the frame of reference in which they make their measurement.

The light clock analysis is appropriate to compare the proper time on the clock in the moving frame of reference (*observed by* Clare in the examples provided in Section 6.2) and the time measured on a clock in the stationary frame (*with* Clare). The light clock was used as it only depends on light, not some complicated mechanical arrangement that may well include other factors that are altered by relative motion. There was, however, one other condition in this clock analysis—that both Amaya and Clare would agree on the distance, d, between the mirrors. This enabled the two expressions for d to be equated in order to find the relationship between proper time, t_0, and time, t.

The clock was deliberately set up in the spaceship so that this light path, of distance d, was perpendicular (at right angles to) the velocity. Distances in this perpendicular direction are unaffected by motion. Indeed, Einstein showed that while perpendicular distances are unaffected, relative motion affects length only in the direction of travel (refer to Figure 6.3.2).

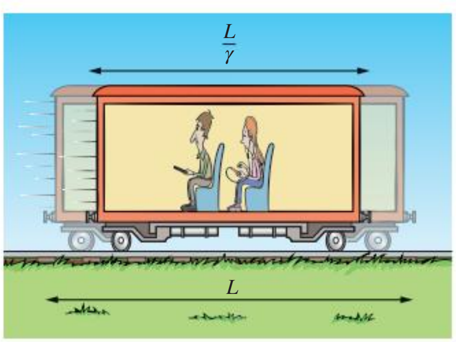


FIGURE 6.3.2 Einstein showed that the length of a moving object is foreshortened by the Lorentz factor, γ. The height and width of the carriage though remain unchanged.

Length contraction

Consider the *Gedanken* situation in which Clare is standing on a train platform while Amaya and Binh pass by at a speed v. Both Clare and Binh want to measure the length of the train platform on which Clare is standing. Using a measuring tape, Clare measures the length of the platform (which is at rest according to her) as L_0, and says that Binh and Amaya cover this distance in a time equal to:

$$t = \frac{L_0}{v}$$

Binh observes the platform passing in a time t_0, as he and Amaya move past the station. The relationship between the time in Binh's frame of reference and the time that Clare measures is:

$$t_0 = \frac{t}{\gamma}$$

$$t_0 = t\sqrt{1 - \frac{v^2}{c^2}}$$

Substituting the first equation into the equation above gives us:

$$t_0 = \frac{L_0}{v}\sqrt{1 - \frac{v^2}{c^2}}$$

Binh sees the platform moving at a velocity of v relative to him, so he can say that the distance from the start to the end of the platform is:

$L = vt_0$

Substituting the previous equation for t_0 into the equation above gives us:

$$L = v \times \frac{L_0}{v}\sqrt{1 - \frac{v^2}{c^2}}$$

This simplifies to:

$$L = L_0\sqrt{1 - \frac{v^2}{c^2}}$$

This is Einstein's **length contraction** equation that incorporates the Lorentz factor. This equation shows that an object with a **proper length** of L_0, when measured at rest, will have a shorter length L, parallel to the motion of its moving frame of reference when measured by an observer that is in a stationary frame of reference. The proper length is contracted by a factor of $\frac{1}{\gamma}$. Length contraction can be represented as:

$$L = \frac{L_0}{\gamma}$$

where $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$

L_0 is the proper length i.e. the length measured at rest, in the stationary frame of reference

L is the length in the moving frame, measured by an observer

Worked example 6.3.1

LENGTH CONTRACTION

Assume *Gedanken* conditions exist in this example. A stationary observer on Earth sees a very fast car travelling by at 2.50×10^8 m s^{-1}. When stationary, the car is 3.00 m long. Calculate the length of the car as seen by the stationary observer. Use $c = 3.00 \times 10^8$ m s^{-1}.	
Thinking	**Working**
Identify the variables: the length measured by the stationary observer is L, the proper length of the car is L_0 and the velocities are v and the constant c.	$L = ?$ $L_0 = 3.00$ m $v = 2.50 \times 10^8$ m s^{-1} $c = 3.00 \times 10^8$ m s^{-1}
Use Einstein's length contraction formula and the Lorentz factor.	$L = \frac{L_0}{\gamma}$ $= L_0\sqrt{1 - \frac{v^2}{c^2}}$
Substitute the values for L_0, v and c into the equation and calculate the answer, L.	$L = 3.00 \times \sqrt{1 - \frac{(2.50 \times 10^8)^2}{(3.00 \times 10^8)^2}}$ $= 3.00 \times 0.553$ $= 1.66$ m

Worked example: Try yourself 6.3.1

LENGTH CONTRACTION

Assume *Gedanken* conditions exist in this example. A stationary observer on Earth sees a very fast scooter travelling by at 2.98×10^8 m s^{-1}. The stationary observer measures the scooter's length as 45.0 cm. Calculate the proper length of the scooter, measured when the scooter is at rest.

LOOKING OUT OF THE WINDOW

So far you have been looking at situations in which objects that are in a moving frame of reference are seen as being shorter in the direction of the motion according to an observer that is in a stationary frame of reference. You can also apply length contraction to the distance that a moving object covers as it travels at very high speed.

Recall that no inertial frame of reference is special. Consider Amaya and Binh in their spacecraft. According to them, they are stationary and it is space itself that rushes by at high speed. As space zooms by Amaya and Binh, they are travelling a proper distance of 384 400 km from the Earth to the Moon. This is the proper length as it is measured by a device that is in the same frame of reference as the Earth and the Moon. As Binh looks out of the window, he sees a much shorter distance to travel.

Worked example 6.3.2

LENGTH CONTRACTION FOR DISTANCE TRAVELLED

Assume *Gedanken* conditions exist in this example. A pilot of a spaceship travelling at 0.997*c* is travelling from Earth to the Moon. The proper distance from the Earth to the Moon is 384 400 km. When the pilot looks out of the window, the distance between the Earth and the Moon looks much less than that. Calculate the distance that the pilot sees.

Thinking	Working
Identify the variables: the length seen by the pilot is L, the proper length of the distance is L_0 and the velocity is v.	$L = ?$ $L_0 = 384\,400$ km $v = 0.997c$ m s^{-1}
Use Einstein's length contraction formula and the Lorentz factor.	$L = \frac{L_0}{\gamma}$ $= L_0\sqrt{1 - \frac{v^2}{c^2}}$
Substitute the values for L_0 and v into the equation. Cancel c and calculate the answer, L.	$L = 384\,400 \times \sqrt{1 - \frac{(0.997c)^2}{c^2}}$ $= 384\,400 \times \sqrt{1 - (0.997)^2}$ $= 384\,400 \times 0.0774$ $= 29\,800$ km

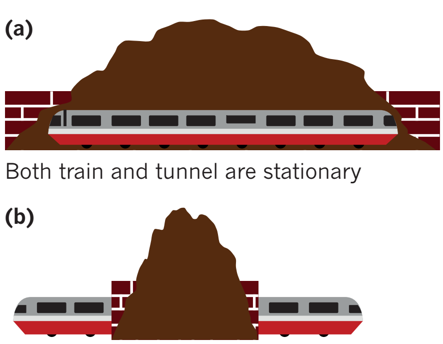

(a) Both train and tunnel are stationary
(b) The tunnel is moving towards the observer in the train

FIGURE 6.3.3 The train both fits in the tunnel and doesn't fit in the tunnel, depending on your frame of reference.

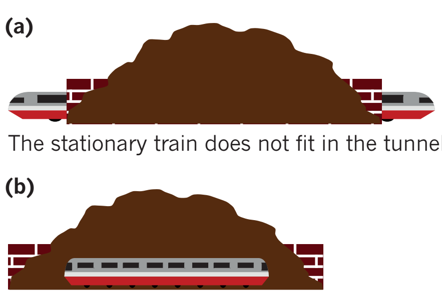

(a) The stationary train does not fit in the tunnel
(b) The train is contracted

FIGURE 6.3.4 The train both doesn't fit in the tunnel and does fit in the tunnel, also depending on your frame of reference.

Worked example: Try yourself 6.3.2

LENGTH CONTRACTION FOR DISTANCE TRAVELLED

Assume *Gedanken* conditions exist in this example. A stationary observer on Earth sees a very fast train approaching a tunnel at a speed of 0.986*c*. The stationary observer measures the tunnel's length as 123 m long. Calculate the length of the tunnel as seen by the train's driver.

The result from Worked example: Try yourself 6.3.2 leads to an interesting phenomenon. If the proper length of the train is 100 m, then the driver could park the train in the 123 m tunnel with 11.5 m of tunnel extending beyond each end of the train. But when the train is moving at 0.986*c*, then according to the train driver the train will not fit in the tunnel. There will be approximately 39.8 m of train extending past each end of the tunnel. This phenomenon is illustrated in Figure 6.3.3.

Similarly, a train that is longer than the tunnel will fit completely inside the tunnel if its length was measured by a stationary observer as it was moving past very quickly. In this scenario, the length of the train would be contracted according to the stationary observer (refer to Figure 6.3.4).

PROPER TIME AND PROPER LENGTH

The time t_0 and the length L_0 are referred to as the proper time and proper length. They are the quantities measured by the observer, who is in the same frame of reference as the event or the object being measured.

Proper time

The proper time is the time between two events that occur at the same point in space. For example, when a light bulb in the train flashes and Amaya measures the time for the flash to reflect off a mirror and return to her, then she has measured the proper time. This is because the stopwatch remained at the point in space inside the frame of reference where the light originated and where it ended up. Proper time is illustrated in Figure 6.3.5.

It is important that a clock isn't moved from one place to another if you want to measure proper time. This is because, as soon as the clock is in motion, the time for that clock slows slightly.

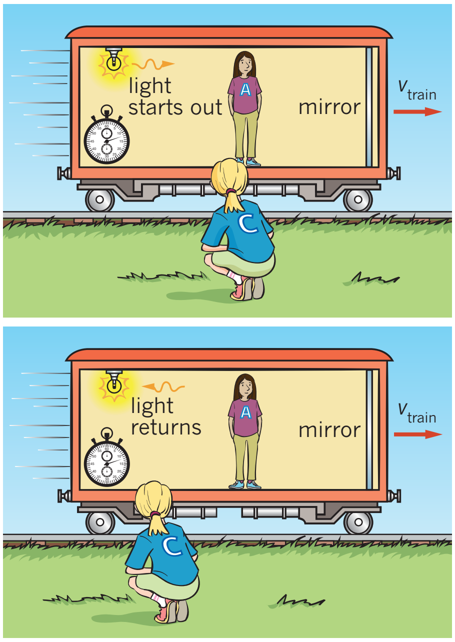


FIGURE 6.3.5 A clock measuring proper time. The clock is positioned at the place where the event started (the light starting out) and is at the same place when the event ends (the light returning).

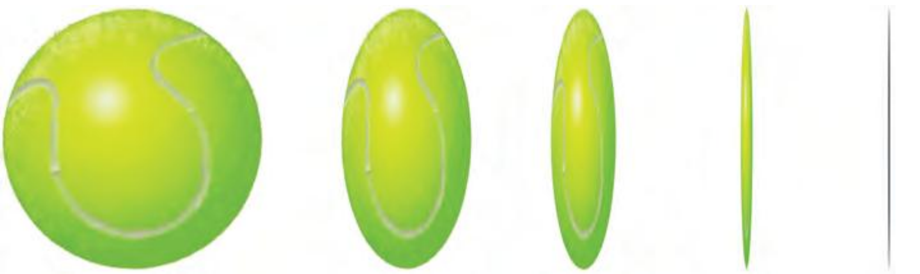
Proper length

The proper length is the distance between two points whose positions are measured by an observer at rest with respect to the two points.

Recall the example of Amaya on a train and Clare on the platform observing the passing train. As Amaya reads her measuring tape at either end of the carriage and is at rest with respect to the train, her measurement of the carriage is the proper length. Clare's measurement of the carriage will be of the contracted length.

Clare, on the other hand, measures the length of the platform as the proper length, while Amaya and Binh see the platform as contracted in length. Remember that length contraction occurs only in the direction of travel, not in any perpendicular direction. To Clare, the carriage will appear shortened, but its width and height (the dimensions of the train perpendicular to the direction of travel) will remain unaltered.

An example of length contraction is shown with a tennis ball in Figure 6.3.6. The length in the direction of the motion is contracted, but the height is not.



$v = 0$	$v = 0.87c$	$v = 0.995c$	$v = 0.999c$	$v \rightarrow c$
$L = L_0$	$L = 0.5L_0$	$L = 0.25L_0$	$L = 0.045L_0$	$L \rightarrow 0$

FIGURE 6.3.6 As the tennis ball moves faster to the right, its length in this dimension is contracted, but its height and depth remain the same.

FINAL THOUGHTS

Length contraction and time dilation are easy to confuse. One way to remember how it works is to think that stationary clocks tick faster and an object is longest when viewed from its own frame of reference. When viewed from a frame of reference where objects are seen to be moving, they appear shorter and their clocks tick slower. All lengths and all clocks seem normal when viewed from within their own frame of reference.

6.3 Review

SUMMARY

- The theory of special relativity states that time and space are related. Motion affects space in the direction of travel.
- A moving object will appear shorter, or appear to travel less distance, by the inverse of the Lorentz factor, γ. Einstein's length contraction equation is given by:
  $L = \frac{L_0}{\gamma}$, where L_0 is the proper length in the stationary frame, L is the contracted length as seen in the moving frame and γ is the Lorentz factor:
  $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$
- The proper time, t_0, is the time measured by an observer at the same point in an inertial frame of reference.
- The proper length, L_0, is the length measured by an observer at rest with respect to the object being measured.

KEY QUESTIONS

For the following questions, assume Gedanken conditions exist and let $c = 3 \times 10^8$ m s^{-1} unless stated otherwise.

1 To what does the term 'proper length', L_0, refer?

2 If you are standing on Earth and observe a speeding rocket ship, what do you notice about its dimensions? Select from the following:
- **A** its length (in the direction of travel) is shorter than normal
- **B** its length (in the direction of travel) is longer than normal
- **C** its height (at right angles to the direction of travel) is shorter than normal
- **D** its width (at right angles to the direction of travel) is shorter than normal

3 An observer is standing on a train platform as a very fast train passes by at a speed of 1.75×10^8 m s^{-1}. The observer notices that a passenger is holding a metre rule in line with the direction that the train is moving. Calculate the length of the metre rule that the stationary observer sees.

4 An observer standing on a comet is watching as a satellite approaches at a speed of 2.30×10^8 m s^{-1}. The observer knows that the proper length of the satellite in the direction of its motion is 5.25 m. Calculate the length of the satellite that the observer sees as it passes.

5 A builder makes a mistake and builds a garage too short for the owner's car to fit in. The proper length of the garage is 1.50 m and the proper length of the car is 3.50 m. The builder suggests that if the owner drives fast enough, the builder could stand by the garage and the car would fit.
- **a** Calculate the speed that the car would need to travel to just fit in the garage when observed by the builder.
- **b** Explain why the car owner would not be happy about the builder's suggestion by calculating the length of the garage as seen by the driver.

6 An observer on a platform measures the time for a train carriage, moving at 0.99c, to pass her by. What time has she measured, t or t_0? Explain.

7 According to a speed (v) versus distance travelled (L) graph, which of the following is true?
- **A** At the maximum speed, the distance travelled is the largest.
- **B** Velocity and distance travelled are directly proportional variables.
- **C** At values close to the speed of light, the distance travelled is near to zero.
- **D** None of the above.

8 Emily is standing by the side of the track, watching Dan run in an 800 m race.
- **a** At what speed must Dan run in order for the race to be only 400 m long in his frame of reference?
- **b** Emily notices that Dan is thinner than he normally is, but just as wide and just as tall. Calculate the fraction of Dan's thickness while he is running to his normal thickness while standing still.

9 A jet plane zooms past an observer standing on the ground at a speed of 660 m s^{-1}. If the length of the jet is 23.5 m when parked on the tarmac, calculate the length that the observer sees the jet.

10 An astronaut in her spaceship is speeding at 0.900c to the Moon. She is holding, in the direction that the spaceship is moving, a fishing rod that is 2.75 m long.
- **a** Determine the length of the rod as observed by an astronaut in the International Space Station.
- **b** What is the length of the fishing rod as observed by the astronaut in the spaceship?

Chapter review

06

KEY TERMS

aether
classical physics
frame of reference
Gedanken
inertial frame of reference
length contraction
Lorentz factor
medium
paradox
postulate
proper length
proper time
simultaneous
spacetime
time dilation

1 Prove that for an object travelling at any possible velocity, the value of the term below must be less than 1.

$$\sqrt{1 - \frac{v^2}{c^2}}$$

2 One of the fastest objects ever made on Earth was the Galileo Probe which, as a result of Jupiter's huge gravity, entered its atmosphere in 1995 at a speed of nearly 50 000 m s^{-1}. Give an estimate of the Lorentz factor for the probe to nine decimal places. (You may use the expression $\gamma \approx 1 + \frac{v^2}{2c^2}$.)

3 In 1905 Einstein put forward two postulates. Which two of the following best summarise them?
- **A** All observers will find the speed of light to be the same.
- **B** In the absence of a force, motion continues with constant velocity.
- **C** There is no way to detect an absolute zero of velocity.
- **D** Absolute velocity can only be measured relative to the aether.

4 Whereabouts on the Earth's surface are we closest to an inertial frame of reference?

5 Which of the following is closest to Einstein's first postulate?
- **A** Light always travels at 3×10^8 m s^{-1}.
- **B** There is no way to tell how fast you are going unless you can see what's around you.
- **C** Velocities can only be measured relative to something else.
- **D** Absolute velocity is that measured with respect to the Sun.

6 Very briefly explain why Einstein said that we must use four-dimensional spacetime to describe events that occur in situations where high speeds and large distances are involved.

7 Imagine that Amaya is at the front end of a train carriage moving forward at 10.0 m s^{-1}. She shines a laser towards Binh, who is at the other end of the carriage. Clare is watching all this from the side of the track. At what velocity does Clare see the light travelling?

8 Which one or more of the following conditions is sufficient to ensure that we will measure the proper time between two events? We must:
- **A** be in the same frame of reference
- **B** be in a frame of reference which is travelling at the same velocity
- **C** be stationary
- **D** not be accelerating with respect to the frame of the two events

9 Spaceships A and B leave the Earth and travel towards Vega, both at a speed of 0.9*c*. Observer C back on Earth sees the crews of A and B moving in 'slow motion'. Describe how the crew of A see the crew of B, and how they see C and the Earthlings moving.
- **A** B will appear normal, C will be sped up.
- **B** B will appear normal, C will be slowed down.
- **C** B will appear slowed down, C will be normal.
- **D** B will appear sped up, C will be slowed down.
- **E** None of these.

10 If you were riding in a very smooth, quiet train with the blinds drawn, how could you tell the difference between the train (i) being stopped in the station, (ii) accelerating away from the station, (iii) travelling at a constant speed?

11 You are in a spaceship travelling at very high speed past a new colony on Mars. Do you notice time going slowly for you; for example, do you find your heart rate is slower than normal? Do the people on Mars appear to be moving normally? Explain your answers.

12 An observer sitting in a very fast jet plane is looking out of the window at a clock placed on top of a mountain. The passenger, using the mountain's clock, notes that it takes a goat 20.0 s to run along a rocky slope. If the plane is flying at a speed of 2.00×10^8 m s^{-1}, calculate how much time has passed on the passenger's clock.

13 A spectator is standing next to the pool clock and watching as a swimmer races at a speed of 2.25×10^8 m s^{-1}. The spectator times on the pool clock that the swimmer completes one stroke in 1.50 s.
- **a** Calculate how much time the spectator sees pass on the swimmer's wristwatch.
- **b** Calculate how much time the swimmer sees has passed on the pool clock, during which time her own wristwatch shows that 1.50 s have passed.

Chapter review *continued*

14 In the twin paradox explanation, when can you say the twin that stays at home ages faster than the twin that goes on the journey?

A during the acceleration phase
B during the deceleration phase
C during both the acceleration and deceleration phases
D during the constant velocity portion of the journey

15 **a** At what speed would a rocket ship be going if it is observed it to be half its normal length?

b The rocket ship is then observed to accelerate to a certain speed so that its length halved again. Did that mean that it doubled its speed? To what speed did it accelerate?

16 Binh and Amaya are playing table tennis in their spaceship. They rush past Clare in her space station at a relative speed of 240 000 km s^{-1}. Binh says that after he hits the ball it returns to his bat after 1.00 s. Their table is 3.00 m long in the direction of their spaceship's motion and is 1.00 m high.

a Calculate the time between hits, as measured by Clare.

b Calculate the length and height of the table, as measured by Clare.

17 Star Xquar is at a distance of 5 light-years from Earth. Space adventurer Raqu heads from Earth towards Xquar at a speed of 0.9*c*.

a For those watching from Earth, how long will it take for Raqu to reach Xquar?

b From Raqu's point of view how long will it take her to reach Xquar?

c Explain why it is that, although Raqu knew that Xquar was 5 light-years from Earth, and that she was to travel at 0.9*c*, it took much less time than might be expected from these figures.

18 The space shuttle travelled at close to 8000 m s^{-1}. Imagine that as it travels east–west it is to take a photograph of Australia, which is close to 4000 km wide. Because of its speed, the space camera will see everything on Earth slightly contracted.

a About how much less than 4000 km wide will Australia appear to be in this photograph?

b Will the north–south dimension of Australia be smaller as well?

19 Imagine that as we watch a traveller from Earth to the star Vega travelling at 99.5% of the speed of light, we will see that their clocks slow down by a factor of about 10 times.

a Explain how this factor of 10 was arrived at.

b Does this mean that they experience this slowing down of time?

c Vega is about 25 light-years from Earth, so in our frame of reference it takes light from Vega 25 years to reach us. How long will it take our space traveller to reach Vega?

d How long will the traveller find that it takes to travel to Vega?

e Does your answer to part d imply that they were able to get to Vega in less time than light? Explain your answer.

20 Muons are high-speed particles that are created some 15 km above the Earth's surface. Classical physics dictates that due to their short lifespans, muons should not ever reach the Earth's surface even though they travel at incredible speeds (approx. 0.992*c*). However, they do. Explain how this is possible, referring to each of the frames of reference of an observer on Earth and the muon itself.

CHAPTER 07
The relationship between force, energy and mass

In 1687 Isaac Newton published *Principia,* in which he outlined the connection between force and motion of bodies with mass. In 1905 Albert Einstein produced an important refinement. While Newton viewed mass, space and time all as absolutes, Einstein postulated that they only appear to be so in the limited contexts covered by classical physics.

As shown in Chapter 6, distance and time are relative concepts. In this chapter it will be revealed that mass, too, is not the absolute that Newton believed it to be. While chemists had long believed that mass was conserved, and in the 19th century physicists had come to believe that energy was conserved, Einstein showed that it was in fact 'mass–energy' that was conserved. Energy can be converted to mass and mass to energy.

Key knowledge

By the end of this chapter, you will have covered material from the study of forces, energy and mass, and you will be able to:

- investigate and analyse theoretically and practically impulse in an isolated system for collisions between objects moving in a straight line: $F\Delta t = m\Delta v$
- investigate and apply theoretically and practically the concept of work done by a constant force using:
  - work done = constant force × distance moved in direction of net force
  - work done = area under force–distance graph
- analyse transformations of energy between kinetic energy, strain potential energy, gravitational potential energy and energy dissipated to the environment (considered as a combination of heat, sound and deformation of material):
  - kinetic energy at low speeds: $E_k = \frac{1}{2}mv^2$; elastic and inelastic collisions with reference to conservation of kinetic energy
  - strain potential energy: area under force–distance graph including ideal springs obeying Hooke's Law: $E_s = \frac{1}{2}k\Delta x^2$
  - gravitational potential energy: $E_g = mg\Delta h$ or from area under a force–distance graph and area under a field–distance graph multiplied by mass
- interpret Einstein's prediction by showing that the total 'mass–energy' of an object is given by: $E_{tot} = E_k + E_0 = \gamma mc^2$ where $E_0 = mc^2$, and where kinetic energy can be calculated by: $E_k = (\gamma - 1)mc^2$
- describe how matter is converted to energy by nuclear fusion in the Sun, which leads to its mass decreasing and the emission of electromagnetic radiation.

7.1 Impulse

When you crash your Ferrari, it is not only how fast you are going that counts, but how quickly you stop!

This is a direct consequence of Newton's second law. If the velocity decreases over a long time, the deceleration is small. Using the mathematical formula $F = ma$, if a is small, you can conclude that the force required to bring the Ferrari to a stop will also be relatively small. If on the other hand, the Ferrari is brought to a halt very rapidly as in Figure 7.1.1, there will be a large deceleration requiring a large force. The force determines the damage. Despite the likelihood of injury caused by a large force that acts for a short time such as in a car crash, a small force acting for a longer time has the same effect on the motion of an object. That is, gradually applying the brakes will also bring the car to rest. One way to quantify the similarity between these situations is to describe the impulse in a **collision**, which considers both the force and the time over which it acts.

FIGURE 7.1.1 Rapid deceleration requires a large force and often results in damage and injury.

CHANGE IN MOMENTUM

Newton's original formulation of his second law was not expressed in terms of acceleration. Rather, he spoke of the '*motion*' of an object that would be altered when a force acted on the body over a time interval. You would say that the **momentum** of the object changes when a resultant force acts on it. This is completely equivalent to the more familiar $F = ma$ formulation of Newton's second law, as shown below.

Consider a body of mass m, with a resultant force F acting on it for a time, Δt. The mass will accelerate as described by Newton's second law:

$F = ma$

$\therefore F = \frac{m\Delta v}{\Delta t}$, using the definition of acceleration.

Rearranging, we can write $F\Delta t = m\Delta v = \Delta p$.

The term $m\Delta v$ is the change in momentum of the object. The force involved in a collision can change in value during the collision, so the average force, F_{ave}, is used. The average force acting on the object for a time Δt causes a change in the momentum.

The term $F_{ave}\Delta t$ is called the impulse of the resultant force and is equal to the change in momentum of the object.

The Ferrari coming to a rapid stop in Figure 7.1.1 loses all of its initial momentum, and a considerable force acts over a very short time interval. The same momentum loss could be achieved much more sedately using the brakes over a longer time interval. The force required would be a lot less, and there would be no damage.

It is important to note that **impulse** is a vector quantity, where the direction is the same as that of the average force or of the change in momentum/velocity.

PHYSICSFILE

Momentum units

Since impulse can be expressed in terms of a momentum change, the units for momentum (kg m s^{-1}) and impulse (N s) must be equivalent. This can be shown by using Newton's second law.

Given that 1 N = 1 kg m s^{-2} (from $F = ma$), it follows that 1 N s = 1 kg m s^{-2} × s

i.e. 1 N s = 1 kg m s^{-1}

Even though the units are equivalent, they should be used with the appropriate quantities as a reminder of the quantity that is being dealt with.

The newton second (N s) is the product of a force and a time interval and so should be used with impulse.

The kilogram metre per second (kg m s^{-1}) is the product of a mass and a velocity and so should be used with momentum.

Even so, it is not uncommon to see newton seconds used as the unit of momentum, nor is this incorrect.

Worked example 7.1.1

CALCULATING THE IMPULSE AND AVERAGE FORCE

Calculate the impulse of a tree on a 1485 kg sports car if the vehicle is travelling at 93.0 km h^{-1} in a northerly direction when the driver loses control of the vehicle on an icy road and the car comes to rest against the tree.

If the car crumples in 40.0 ms, find the average force exerted by the tree on the car.

Thinking	Working
Convert the speed to m s^{-1}.	93.0 km h^{-1} = $\frac{93.0}{3.6}$ m s^{-1} = 25.8 m s^{-1}
Calculate the change in momentum. The negative sign indicates that the change in momentum, and therefore the impulse, is in the direction opposite to the initial momentum (north is positive), as would be expected.	$\Delta p = m(v - u)$ $= 1485(0 - 25.8)$ $= -3.84 \times 10^4$ kg m s^{-1}
The impulse is equal to the change in momentum.	Impulse = 3.84×10^4 kg m s^{-1} south
Transpose $\Delta p = F_{ave}\Delta t$ to find the force.	$F_{ave} = \frac{\Delta p}{\Delta t}$ $= \frac{-3.84 \times 10^4}{40.0 \times 10^{-3}}$ $= -9.60 \times 10^5$ N $= 9.60 \times 10^5$ N south

Worked example: Try yourself 7.1.1

CALCULATING THE IMPULSE AND AVERAGE FORCE

Prior to the accident, the driver had stopped to refuel. Calculate the impulse of the braking system on the 1485 kg car if the vehicle was travelling at 95.5 km h^{-1} in a north-easterly direction and the driver took 12.5 s to come to a halt. Also find the average braking force.

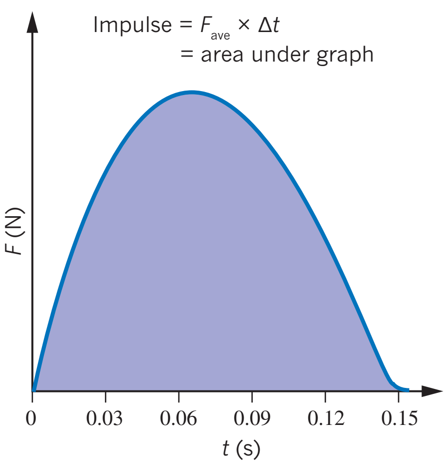


FIGURE 7.1.2 The force changes with time as the racquet strikes the ball.

FORCE VERSUS TIME GRAPHS

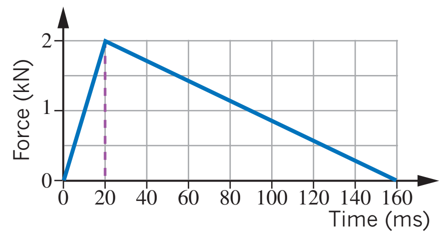
In many practical applications the force applied to an object is not constant. For example, when a tennis player hits a ball, the initial force exerted by the racquet is relatively small. As the strings stretch and the ball deforms, this force builds up to a maximum value before decreasing again as the ball rebounds from the racquet. If the force is measured over time, this data is often represented graphically (see Figure 7.1.2).

The impulse of the ball, or the change in momentum, may be found from the product $F_{ave}\Delta t$. This is simply the area under the force–time graph.

Worked example 7.1.2

RUNNING SHOES

A running-shoe company plots the following force–time graph for a running shoe. Use the data to calculate the magnitude of the impulse.

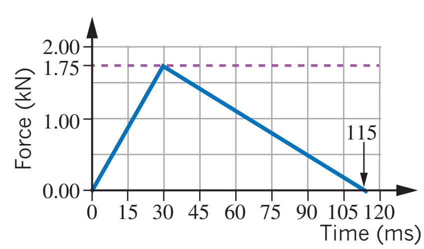


Thinking	Working
Recall that impulse = $F_{ave}\Delta t$. When the force is not constant, this is the area under the force–time graph.	impulse = $\frac{1}{2}$ base × height = $\frac{1}{2} \times 160 \times 10^{-3} \times 2 \times 10^{3}$ = 160 N s

Worked example: Try yourself 7.1.2

RUNNING SHOES

A running-shoe company plots the following force–time graph for an alternative design intended to reduce the peak force on the heel. Calculate the magnitude of the impulse.



Note that in the worked examples above, the impulse is almost the same but the momentum has been transferred over a longer time, resulting in a lower maximum force.

APPLICATIONS

The connection between impulse, force and collision duration is useful for the analysis of collisions. When a vehicle collides with another object and comes to rest, the vehicle and occupants undergo a rapid deceleration. The impulse depends on the initial speed of the vehicle, and on its mass as well.

Since impulse = Δp = $F_{ave}\Delta t$, a large force is exerted to bring the vehicle to rest over a very short time. Extending the time taken for a vehicle to stop reduces the force exerted. Examples of increased stopping times in different activities are shown in Figure 7.1.3.

Worked example 7.1.3

BRAKING FORCE

Calculate the magnitude of the average force exerted by the brakes of a 2500 kg truck to bring the vehicle to rest in 12.0 s, if it were travelling at 30.0 m s^{-1} before the brakes were applied.	
Thinking	**Working**
The change in momentum, $\Delta p = F_{ave}\Delta t$. Calculate the change in momentum. The negative sign indicates that the change in momentum, and therefore the braking force, is in the direction opposite to the initial momentum, as would be expected.	$\Delta p = m\,(v - u)$ $= 2500\,(0 - 30.0)$ $= -7.5 \times 10^4$ kg m s^{-1}
Transpose $\Delta p = F_{ave}\,\Delta t$ to find the force. The sign of the momentum can be ignored since you are finding the magnitude of the force.	$F_{ave} = \frac{\Delta p}{\Delta t}$ $= \frac{7.5 \times 10^4}{12}$ $= 6.25 \times 10^3$ N

Worked example: Try yourself 7.1.3

BRAKING FORCE

The same 2500 kg truck travelling at 30.0 m s^{-1} needs to stop in 1.5 s because a vehicle up ahead stops suddenly. Calculate the magnitude of the braking force required to stop the truck.

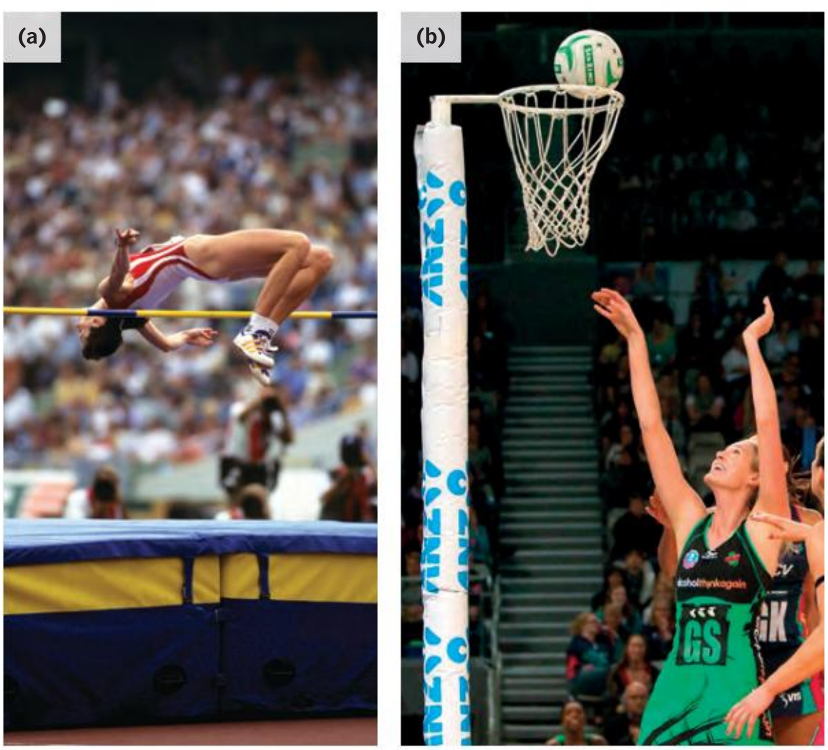


FIGURE 7.1.3 (a) The landing mat extends the time over which the athlete comes to rest, reducing the size of the stopping force. If the high jumper missed the mat and landed on the ground, the force would be larger, but their momentum change would be the same. (b) Thick padding around the goal post extends the time over which a player comes to rest in a collision, thereby reducing the size of the stopping force. (c) Wicketkeepers allow their hands to 'give' when keeping to a bowler. This extends the ball's stopping time and reduces the stopping force.

Safety features such as crumple zones and airbags (see Figure 7.1.4) are aimed at extending Δt, which reduces the force on the occupants of the vehicle, potentially saving lives and preventing injuries.

FIGURE 7.1.4 Airbags reduce the force on passengers by extending the time that it takes for them to stop.

Figure 7.1.5 shows a force–time graph for a collision when an airbag is inflated compared with one where there is no airbag. The change in momentum, or impulse, of the passenger is the same in each case. Thus the area under each curve should be equal. Note, however, that both the peak force and the average force are significantly higher where there is no airbag. The broader peak for the airbag indicates that the passenger is losing their momentum over a longer time, requiring a lower force.

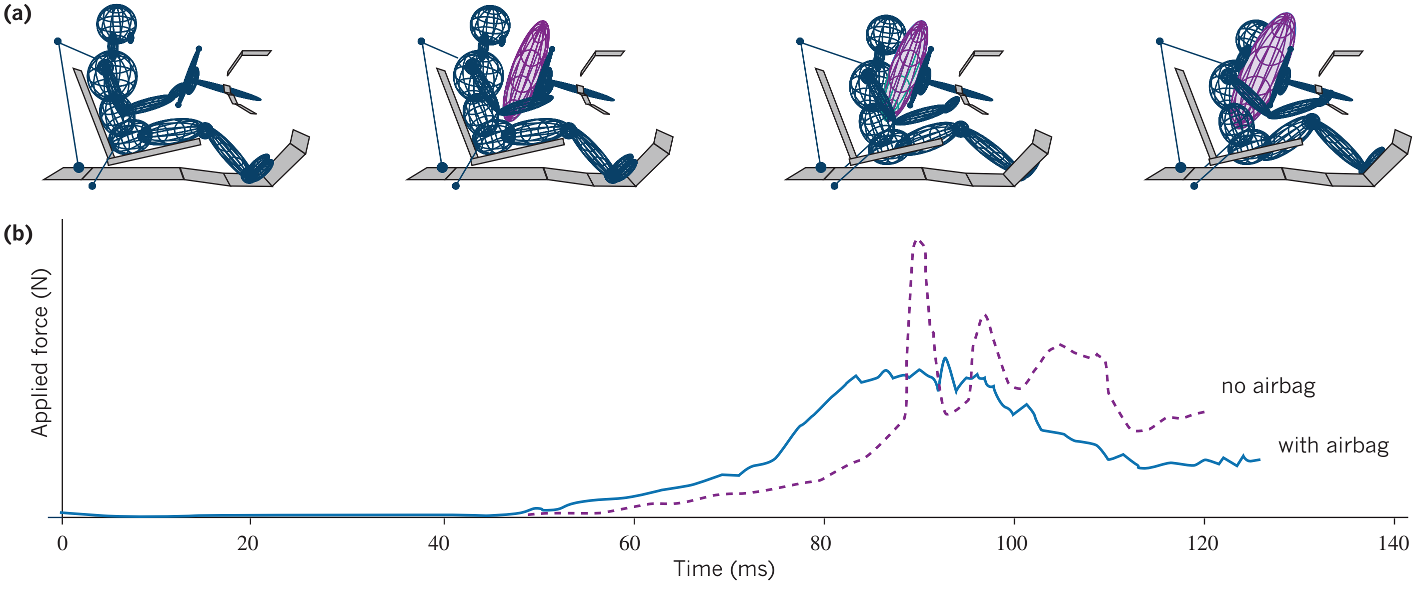


FIGURE 7.1.5 Graph illustrating the difference in the force on a passenger over time when an airbag is inflated in a collision (solid line), and when no airbag is present (dotted line).

PHYSICS IN ACTION

Car safety and crumple zones

Worldwide, car accidents are responsible for over 2 million deaths each decade. Many times this number of people are injured. One way of reducing the road toll is to design safer vehicles. Modern cars employ a variety of safety features that help to improve the occupants' survival chances in an accident. Some of these safety features are the antilock braking system (ABS), electronic stability control (ESC), inertia reel seatbelts, variable-ratio-response steering systems, collapsible steering columns, head rests, shatterproof windscreen glass, padded dashboards, front and side air bags, front and rear crumple zones and a rigid passenger compartment.

Some prototype cars today are equipped with collision avoidance systems. These have laser or infrared sensors that advise the driver of hazardous situations, and even take over the driving of the car in order to avoid accidents! Consider the example of the driver of a car that crashes into a tree at 60 km h^{-1} (16.7 m s^{-1}). If the driver has a mass of 90 kg, then the momentum of the driver is:

$p = m \times v = 90 \times 16.7 = 1500$ kg m s^{-1}

As a result of this collision, the driver will lose all of this momentum as the car comes suddenly to a stop. The impulse experienced by the driver is the same whether the stop is sudden or gradual. In either a sudden or gradual stop, the impulse is –1500 N s. So the idea of safety features such as inertia reel seatbelts, collapsible steering columns, padded dashboards, air bags and crumple zones is not to reduce the size of the impulse, but to reduce the size of the forces that act to bring the driver to a stop. Automotive engineers strive to achieve this by extending the time over which the driver loses momentum.

Crumple zones

A popular misconception among motorists is that cars would be much safer if they were sturdier and more rigid. Drivers often complain that cars seem to collapse too easily during collisions, and that it would be better if cars were structurally stronger—more like an army tank. In fact, cars are specifically designed to crumple to some extent (see Figure 7.1.6). This makes them safer and actually reduces the seriousness of injuries suffered in car accidents.

Army tanks are designed to be extremely sturdy and rigid vehicles. They are able to withstand the effect of collisions without suffering serious structural damage. If a tank travelling at 50 km h^{-1} crashed into a solid obstacle, the tank would be relatively undamaged. However, its occupants would very likely be killed. This is because the tank has no 'give' in its structure and so the tank and its occupants would stop in an extremely short time interval. The occupants would lose all of their momentum in an instant, which means that the forces acting on them would necessarily be extremely large. These large forces would cause the occupants of the tank to sustain very serious injuries, even if they were wearing seatbelts.

FIGURE 7.1.6 Cars are designed with weak points in their chassis at the front and rear that enable them to crumple in the event of a collision. This extends the time over which the cars come to rest and so reduces the size of the forces acting on the occupants.

Cars today have strong and rigid passenger compartments; however, they are also designed with non-rigid sections such as bonnets and boots that crumple if the cars are struck from the front or rear (see Figure 7.1.7). The chassis contains parts that have grooves or beads cast into them. In a collision, these beads act as weak points, causing them to crumple in a concertina shape.

FIGURE 7.1.7 The Australian New Car Assessment Program (ANCAP) assesses the crashworthiness of new cars. This car has just crashed at 50 km h^{-1} into a 5-tonne concrete block. The crumpling effect can clearly be seen.

This 'concertina' effect allows the front or rear of the car to crumple, extending the time interval over which the car and its occupants come to a stop. This stopping time is typically longer than 0.1 s in a 50 km h^{-1} crash. Because the occupants' momentum is lost more gradually, the peak forces that act on them are smaller and so the chances of injury are reduced.

7.1 Review

SUMMARY

- When a force is exerted on an object for a time interval Δt, it brings about a change of momentum Δp by changing the velocity of the object:
  $F_{ave}\Delta t = m\Delta v = \Delta p$
- The impulse is the change of momentum of an object.
- Impulse can be calculated from the area under a force–time graph.
- In many practical applications measures are taken to increase the time of the interaction in order to reduce the maximum force.

KEY QUESTIONS

1 Three balls of identical mass are thrown against a surface at the same speed.
Ball A stops on impact.
Ball B rebounds with 75% of its initial speed.
Ball C rebounds with 50% of its initial speed.
Order the balls in terms of their change in momentum, from least to greatest.

2 When a tennis player serves, she hits the 57 g tennis ball at the top if its flight so that the ball is momentarily stationary, and leaves the racquet at 144 km h^{-1}. If the ball and racquet are in contact for 0.060 s, calculate the magnitude of the average force exerted by the racquet on the ball.

3 A 160 g cricket ball flies past the wicket at 155 km h^{-1} and is stopped by the wicket keeper. Calculate the magnitude of the impulse delivered by the ball to the wicket keeper.

4 A child wearing a backpack jumps from a tree, landing on her feet on the ground. Select the correct statements about factors that will influence the force on her knees and ankles when she lands. More than one correct answer is possible.

A Wearing good runners will reduce the force, while being barefoot will increase the force.
B Bending her knees as she lands will increase the force.
C Landing on concrete will increase the force compared to landing on grass.
D Jumping from a lower branch will decrease the force.
E Leaving the backpack in the tree will make no difference to the force.

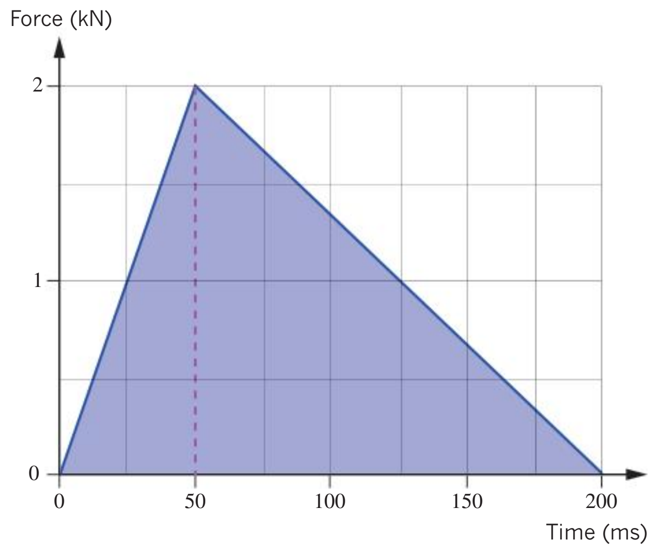
5 A basketball of mass 0.625 kg is bounced against the court at a speed of 32.0 m s^{-1} and it rebounds at 24.5 m s^{-1}. Calculate the average force exerted by the court on the ball if the interaction lasts 16.5 ms.

6 **a** Calculate the momentum of a 100 tonne train travelling at 50.0 km h^{-1}.
b Calculate the magnitude of the impulse if the train were to collide with a 5.00 tonne truck at a level crossing, and push the truck for 15 m before coming to rest.

7 Jacinta does a physics experiment in which she inflates two basketballs to different pressures, and on dropping them from the same height, finds that the ball with the highest pressure bounces higher. She argues that if each of the two balls were thrown with the same speed at a player, the ball with higher pressure would hurt more. Her practical partner Sarah disagrees. She says the balls have essentially the same mass, the same speed, and so they will not be any different when they hit the player. Who is correct? Use the concepts from this chapter to justify your answer.

The following information relates to questions 8 and 9.

The figure below shows a schematic representation of the force exerted by an athlete's foot over the 200 ms that his foot is in contact with the ground.



8 Calculate the magnitude of the impulse of the athlete on the ground.

9 Calculate the magnitude of the average force exerted by his foot over the duration of the contact.

10 A tennis ball of mass 57.5 g is tested for compliance with tennis regulations by being dropped from a height of 251 cm onto concrete. A bounce height of 146 cm is deemed acceptable. Find the magnitude of the average force on the ball if it is in contact with the floor for 0.0550 s.

7.2 Work done

In everyday language, the concept of work is associated with effort and putting energy into something, whether it be in your studies, sports, or even an important relationship. Although the word '*work*' has a much more specific meaning in Physics, it is still connected with energy.

As you saw in the previous section, when a force acts on an object over a time interval, the object accelerates and its momentum changes. When the force causes a displacement in the direction of the object, the energy of the object changes, and we say that work has been done. The weightlifter in Figure 7.2.1 does work by exerting a force, causing the barbell she is lifting to undergo a displacement. The gravitational potential energy of the barbell is increased, and the store of chemical energy in the muscles of the weightlifter is decreased.

FIGURE 7.2.1 If this weightlifter lifts a 100 kg barbell, she has to exert a force of 980 N to oppose the force of gravity. When she lifts the barbell through a height of 0.5 m, she does 490 J of work on the barbell, and increases the gravitational potential energy of the barbell by the same amount.

CALCULATING WORK

Work is the transfer of energy from one object to another and/or the transformation of energy from one form to another. A force does work on an object when it acts on a body causing a displacement in the direction of the force. Where the force is constant, the *work done by* the force is: $W = Fs$.

If the force is applied at an angle to the displacement, only the component of the force in the direction of the displacement contributes to the work done. That is, if the force and displacement vectors are at an angle θ with respect to each other, then $F \cos \theta$ is the component of force that does work:

$$W = Fs \cos \theta$$

where W is the work done by the force (J)

F is the magnitude of the constant force (N)

s is the displacement (m)

θ is the angle between the force vector and the displacement vector

While both force and displacement are vectors, work and energy are scalar quantities that are measured in joules (J).

PHYSICSFILE

The joule

The unit for work is the newton metre (N m), which is called a joule (J) in honour of James Joule, an English physicist who did pioneering work on energy in the 19th century. All forms of energy are measured in joules. The units for gravitational potential energy, for example, can be shown to be equivalent to joules.

$E_g = mg\Delta h$

Units for E_g = kg m s^{-2} m

= N m

= J

You might like to try this for kinetic energy.

From Unit 2 Physics, you will remember that the unit for torque is also the newton metre; however in that case it is not the same as a joule. The reason for this is that, for torque, the force and the lever arm are perpendicular to each other whereas for work, the force and the displacement are parallel to each other.

When you want to find the work done on an object, it is the net force that needs to be used. For instance, if a man pushes a heavy couch across a carpeted floor, the work done on the couch depends on the force applied by the man, less the frictional force which opposes the motion:

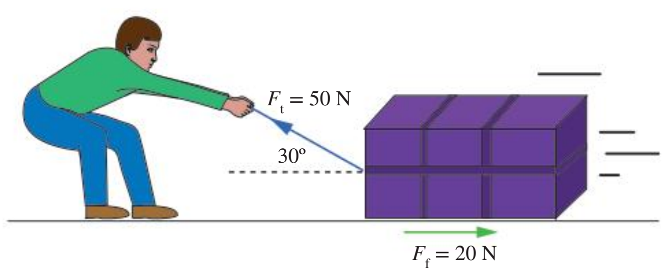
$W = \Delta E = F_{net}\, s$

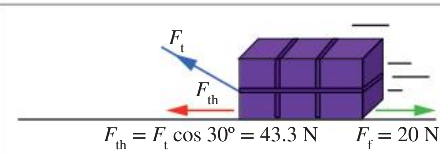
The energy, ΔE, gained by the couch depends on the net force acting on it.

Worked example 7.2.1

WHEN THE FORCE APPLIED IS AT AN ANGLE TO THE DISPLACEMENT

A rope that is at 30.0° to the horizontal is used to pull a 10.0 kg crate across a rough floor. The crate is initially at rest and is dragged for a distance of 4.00 m. The tension, F_t, in the rope is 50.0 N and the frictional force, F_f, opposing the motion is 20.0 N.



a Determine the work done on the crate by the person pulling the rope.	
Thinking	**Working**
Draw a diagram of the forces in action.	
Find the component of the tension in the rope, F_t, that is in the direction of the displacement, i.e. F_{th}.	$F_{th} = 50 \times \cos 30° = 43.3$ N
Find the work done by the person (this includes work done on the crate, and work done against friction).	$W = F_{th}s$ $= 43.3 \times 4.00$ $= 173.2$ J

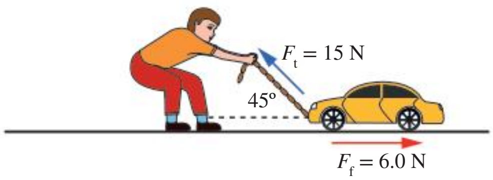
b Calculate the work done on the crate.	
Thinking	**Working**
The work done on the crate is the net force multiplied by the displacement. (This is the increase in the kinetic energy of the crate.)	$W = F_{net}s = (F_{th} - F_f)s$ $= (43.3 - 20) \times 4.00$ $= 93.2$ J

c Calculate the energy transformed to heat and sound due to the frictional force.	
Thinking	**Working**
Energy transformed to heat and sound due to the frictional force is the difference between the work done by the person and the energy gained by the crate.	Energy = 173.2 − 93.2 = 80.0 J
This is equal to the work done against friction, which could also be calculated from the frictional force.	$W_f = F_f\, s$ $= 20.0 \times 4.00$ $= 80.0$ J

Worked example: Try yourself 7.2.1

WHEN THE FORCE APPLIED IS AT AN ANGLE TO THE DISPLACEMENT

A boy drives a toy car by pulling on a cord that is attached to the cart at 45° to the horizontal. The boy applies a force of 15.0 N and pulls the car for 10.0 m down a pathway against a frictional force of 6.0 N.



a Determine the work done on the car by the boy pulling on the cord.

b Calculate the work done on the toy car.

c Calculate the energy transformed to heat and sound due to the frictional force.

When a force performs no work

It is important to remember that work is only done when a force, or a component of force, is applied in the direction of displacement. Hence it is possible to exert a force and feel very tired without doing work. This would mean no energy has been transferred. For example, if you hold a heavy object with your arms out in front of you, you will get tired very quickly but you are not doing any work on the object.

Similarly, an object moving in a circular path in a horizontal plane is constantly accelerated by the centripetal force. Because this force is perpendicular to the displacement at each instant, the force does no work, and no energy is transferred to the object. It does not get faster or slower, it only changes direction as shown in Figure 7.2.2.

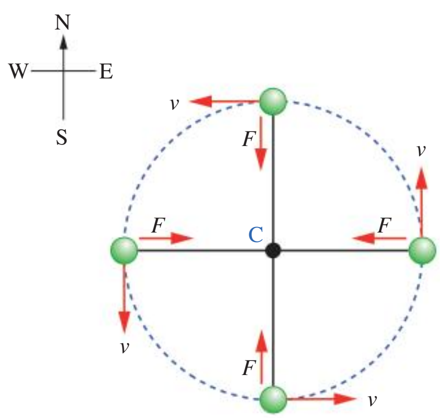


FIGURE 7.2.2 A body moving in a circular path has a force directed towards the centre. The displacement is in the direction of the velocity. There is therefore no force in the direction of the displacement.

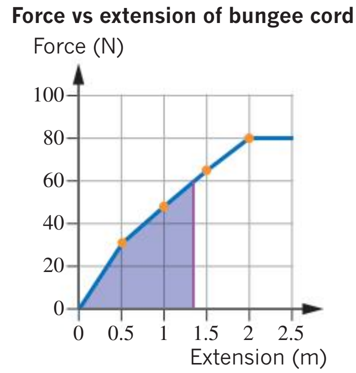
Force–distance graphs

When the force is constant the work done is easily calculated, but in many practical applications the net force is not constant. Where the force–distance relationship is represented graphically, the work done is the area under the force–distance graph. This principle is very similar to the way in which the impulse can be calculated from the area under a force–time graph, as described in the previous section. However, it is important not to confuse these two quantities.

Worked example 7.2.2

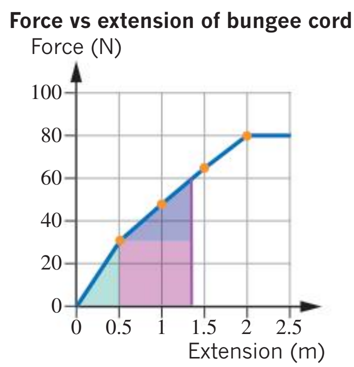
CALCULATING WORK FROM A FORCE–DISTANCE GRAPH

The force required to stretch a piece of bungee cord was recorded in the graph below. Calculate the work done when a 60 N force is applied to the cord.



Thinking

The work done is the area under the force–distance curve. This may be found by calculation, or by counting squares. In this case it is best to divide the area into triangles and rectangles.



Working

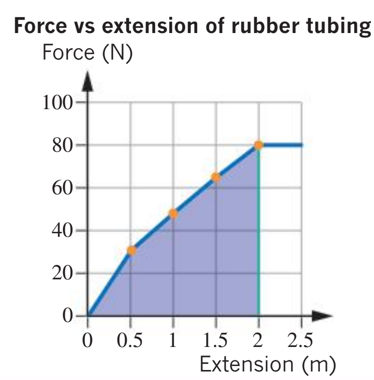
$$\text{Area} = \frac{1}{2} \times 0.50 \times 30 + \frac{1}{2} \times 0.90 \times 30 + 30 \times 0.90$$

Work done = 48 J

Worked example: Try yourself 7.2.2

CALCULATING WORK FROM A FORCE–DISTANCE GRAPH

The force required to elongate a piece of rubber tubing was recorded in the graph below. Calculate the work done when the rubber was stretched by 2.0 m.



If the force–distance curve is not linear, the area can be estimated by counting squares under the curve. It is important to take careful note of the units in order to calculate the work represented by each square.

7.2 Review

SUMMARY

- When a force does work on an object, there is a change in the displacement and the energy of the object.
- Work, W, is a scalar and is measured in joules (J).
- The work done on an object is the net force on the object multiplied by the distance moved in the direction of the force: $W = F_{net}s$.
- When the force is applied to an object at an angle to the displacement, work is only done by the component of the force in the direction of the force: $W = Fs \cos \theta$.
- A centripetal force does no work on an orbiting object, as the force and displacement are perpendicular.
- The work done by a varying force is the area under the force–distance graph.

KEY QUESTIONS

1 Select the scenario in which no work is done.
A Janet stands on a horizontal travelator holding a suitcase above the ground.
B James walks across the airport lounge with his backpack on his back, and climbs a flight of stairs to the boarding gate.
C Jeremy lifts his suitcase and stands holding it for a few minutes.
D Jason wheels his suitcase across the floor.

The following information applies to questions 2–4.

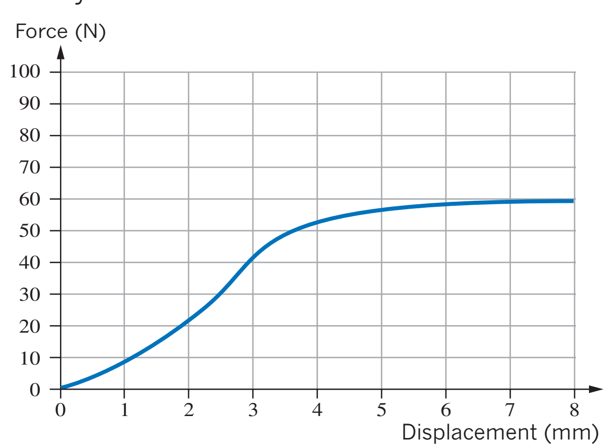
A child uses a leash to drag a reluctant 2.0 kg puppy across a floor. The leash is held at an angle of 60° to the horizontal and the child applies a force of 30 N on the puppy, which is initially at rest. A constant frictional force of 10 N acts on the dog as it is dragged for a distance of 2.4 m.

2 Calculate the work done by the horizontal component of the 30 N force.

3 Calculate the work that the child does in overcoming friction.

4 Calculate the kinetic energy gained by the puppy.

5 The graph below shows the force–distance curve as a sports shoe is compressed during the stride of an athlete. Estimate the work done in compressing the shoe by 7 mm.



6 Earth orbits the Sun and experiences a constant gravitational force. Select the correct statement.
A The work done by the gravitational force in one orbit is the magnitude of the gravitational force multiplied by the circumference of the orbit.
B The gravitational force does not do work on Earth because it is not a contact force.
C The work done by the gravitational force is equal to the kinetic energy of Earth.
D The gravitational force does not do work on Earth because the force and displacement are perpendicular.

7 A weightlifter raises a 150 kg barbell to a height of 1.20 m at constant speed. Calculate the work done by the weightlifter.

8 A proton moving with a velocity of 4.0×10^6 m s^{-1} in a magnetic field of strength 1.7 T experiences a force $F = qvB$ which causes the charge to travel in a circular path of radius 2.5 cm. Calculate the work done by the force in one revolution.

9 In the javelin event the javelin is released at an angle in order to achieve maximum flight distance. An 800 g javelin is released at an angle of 45° at a height of 1.9 m, and at a speed of 108 km h^{-1}. Calculate the work done by the gravitational force on the javelin from its release to the point where it lands on the ground.

10 Krisha pushes a lawnmower at constant speed across 15 m of lawn. She applies a force of 68 N at an angle of 60° to the horizontal. Calculate the work that she does against friction.

7.3 Strain potential energy

FIGURE 7.3.1 The strain potential energy stored in the pole is what allows the pole vaulter to propel herself over the bar.

In everyday life, you frequently encounter situations in which work is done to stretch or compress materials. Think of sports like bungee jumping, pole vaulting (Figure 7.3.1), trampolining and tennis, where the elastic properties of materials are harnessed to generate thrills for spectators and participants. Computer keyboards have tiny springs in the keys, while windup toys, old-fashioned watches, door-closing mechanisms and car suspensions are some of the other devices that use springs.

HOOKE'S LAW

Initially, it is relatively easy to start stretching a spring, but more and more force is required for each incremental amount of extension. This is expressed in Hooke's law:

$F_s = -k\Delta x$

where F_s is the force exerted by the spring (N)

k is the spring constant (N m^{-1})

Δx is the displacement (the extension or compression) of the spring (m)

The force exerted by a spring is directly proportional to, but opposite in direction to, the spring's extension or compression. The spring constant k is a measure of the stiffness of the spring. The behaviour of a spring under force is often illustrated graphically by plotting the force applied versus the extension achieved, as shown in Figure 7.3.2. Notice that a stiffer spring has a larger spring constant, and the spring constant is represented by the gradient (slope) of the graph.

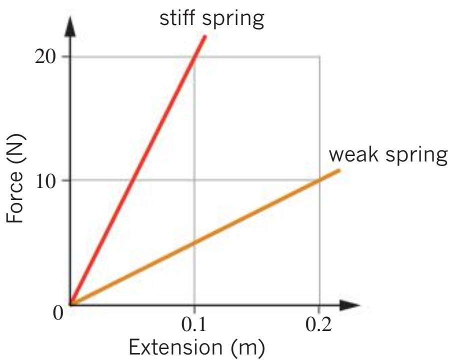


FIGURE 7.3.2 Both springs represented in this graph are ideal. The springs obey Hooke's law because they both have linear graphs, but they have different degrees of stiffness. The stiff spring has a spring constant of 200 N m^{-1}. The spring constant of the other spring is just 50 N m^{-1}. The stiffer spring has the higher gradient (steeper line) on the F–x graph.

PHYSICSFILE

Climbing ropes

The ropes used by rock climbers have elastic properties that can save lives during climbing accidents. Ropes that were used in the 19th century were made of hemp, which is strong but does not stretch a lot. When climbers using these ropes fell, they stopped very abruptly. The resulting large forces acting on the climbers caused many serious injuries. Modern ropes are made of a continuous-drawn nylon fibre core and a protective textile covering. They have a slightly lower spring constant and stretch significantly (up to several metres) when stopping a falling climber. This reduces the stopping force acting on the climber. Ropes with even lower spring constants are suitable for bungee jumping. Rock climbers tend to avoid these ropes—bouncing up and down the rock face is not advisable!

When considering the work done in deforming a spring, the force applied is in the direction of the displacement and hence the negative sign in $F = -k\Delta x$ falls away. The applied force is a linear function of distance and, as discussed in the previous section, when force is not constant the work done by the force may be calculated from the area under the force–distance curve. Hence this technique can be used to determine the work required to extend or compress a spring.

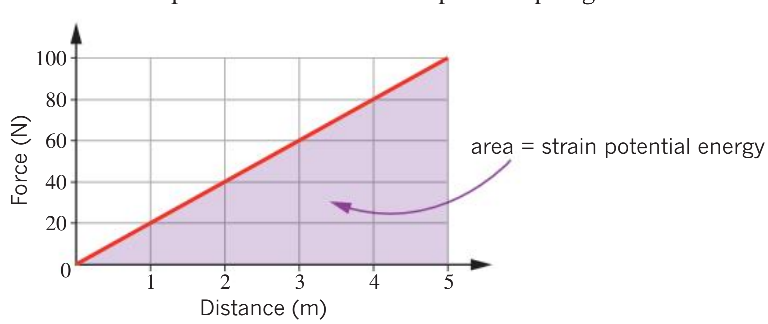


FIGURE 7.3.3 The strain potential energy is calculated by the area under the force–distance curve.

This leads us to an expression for the work done in deforming a spring, and the **strain potential energy** which is then stored in the spring:

$E_s = \frac{1}{2}$ height × base of the shaded triangle in Figure 7.3.3, thus:

$= \frac{1}{2} F \times \Delta x$

$= \frac{1}{2} k\Delta x \times \Delta x$

The strain potential energy, E_s, is calculated using:

$$E_s = \frac{1}{2} k\Delta x^2$$

where k is the spring constant (N m^{-1})

Δx is the spring extension (m)

We call the directly proportional relationship between force and extension **elastic** behaviour. Ideal springs obey Hooke's law, and will return all of the strain potential energy when the applied force is removed.

It is possible to exceed the **elastic limit** of a spring or other elastic material. At this point permanent **deformation** occurs. If the force is increased further, the **breaking point** is reached, at which the material fails. These points are shown in Figure 7.3.4.

While the work done in deforming the spring can still be calculated from the area under the force–distance curve, the energy stored may not all be recoverable as work has been done to permanently change the material.

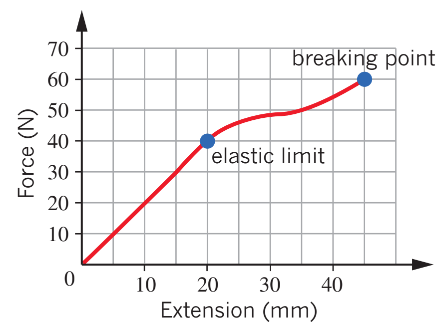
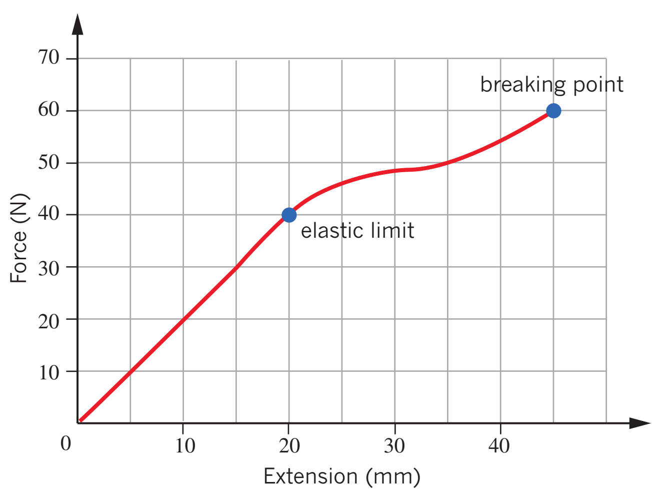


FIGURE 7.3.4 The point at which the force–distance curve deviates from linear behaviour is the elastic limit, where permanent damage is done to the spring.

Worked example 7.3.1

CALCULATING THE SPRING CONSTANT, STRAIN POTENTIAL ENERGY AND WORK

A fine steel wire has the force–extension properties shown in the figure below.



a Calculate the spring constant k for the wire.	
Thinking	**Working**
The spring constant is the gradient of the linear section of the force–extension curve in units N m^{-1}.	$k = \frac{\Delta F}{\Delta x}$ $= \frac{40}{20 \times 10^{-3}}$ $= 2.0 \times 10^{3}$ N m^{-1}

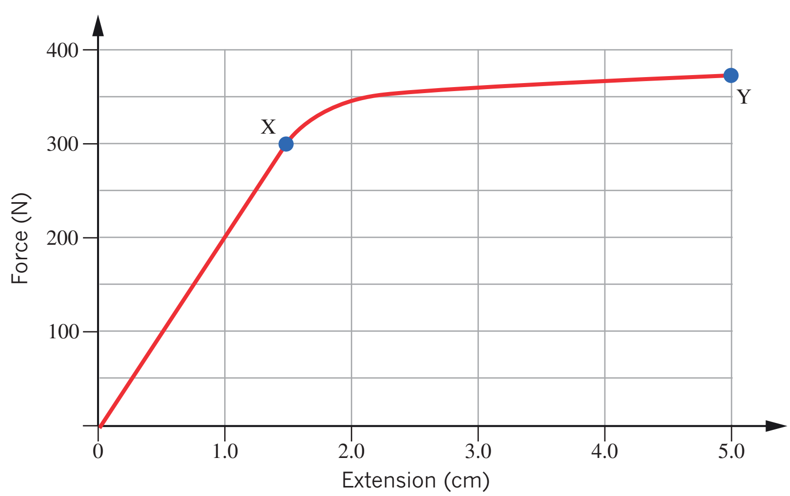
b Calculate the strain potential energy that the wire can store before permanent deformation begins.	
Thinking	**Working**
The strain potential energy is the area under the curve up to the elastic limit.	$E_s = \frac{1}{2}$ height × base of triangle $= \frac{1}{2} \times 4.0 \times 2.0 \times 10^{-3}$ $= 4.0 \times 10^{-3}$ J
This value can also be obtained using the formula for strain potential energy.	$E_s = \frac{1}{2} k\Delta x^2$ $= \frac{1}{2} \times 2.0 \times 10^3 \times (2.0 \times 10^{-3})^2$ $= 4.0 \times 10^{-3}$ J

c Calculate the work done to break the wire.	
Thinking	**Working**
Estimate the number of squares under the curve up to the breaking point.	Number of squares = 33 (approx.)
Calculate the energy per square. The energy per square is given by the area of each square.	Energy for one square = $10 \times 5 \times 10^{-3}$ $= 5.0 \times 10^{-2}$ J
Multiply the number of squares by the energy per square.	Work = energy per square × number of squares $= 5.0 \times 10^{-2} \times 33$ = 1.7 J

Worked example: Try yourself 7.3.1

CALCULATING THE SPRING CONSTANT, STRAIN POTENTIAL ENERGY AND WORK

An alloy sample is tested under tension, giving the graph shown below where X indicates the elastic limit, and Y the breaking point.



a Calculate the spring constant k for the sample.

b Calculate the strain potential energy that the alloy can store before permanent deformation begins.

c Calculate the work done to break the sample.

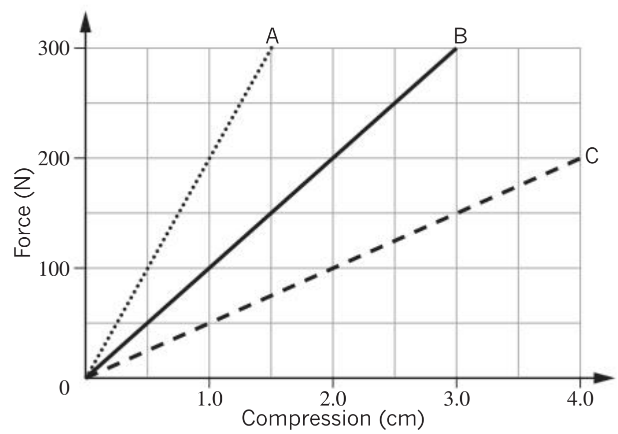
7.3 Review

SUMMARY

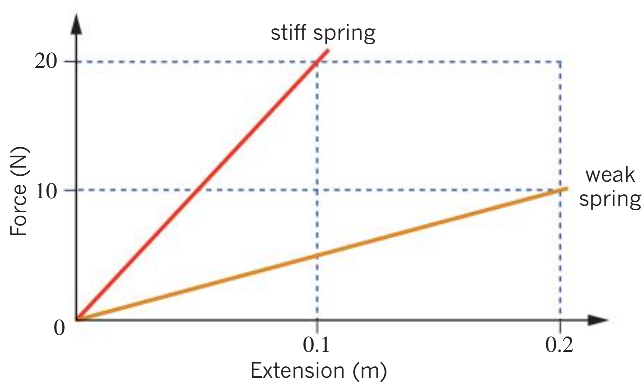
- Hooke's law states that the force exerted *by* a spring is $F_s = -k\Delta x$. The negative sign indicates that the force opposes the displacement.
- k is the spring constant and is measured in N m^{-1}. This may be calculated from the gradient of the linear section of a force–displacement graph.
- The work done to deform an ideal spring is equal to the strain potential energy stored in the spring: $E_s = \frac{1}{2}k\Delta x^2$
- When a material displays elastic behaviour it obeys Hooke's law, and the strain potential energy stored is returned when the force is removed.
- When a material exceeds its elastic limit, permanent deformation occurs and not all of the strain potential energy is returned when the force is removed.

KEY QUESTIONS

1 Rank the springs below in order of increasing stiffness.



The following information relates to questions 2–3.
The graph of stretching force versus extension for two different springs is shown below.



2 Calculate the spring constant for both springs.

3 Find the difference between the strain potential energy stored when each of the springs in the graph above is extended by 20 cm.

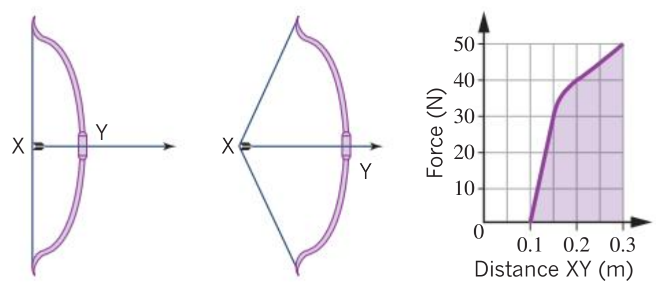
4 A 1.0 m piece of rubber has a spring constant of 50 N m^{-1}. Calculate how much the rubber will stretch if a force of 4.0 N is exerted on it.

The following information relates to questions 5–6.
A toy plane is launched by using a stretched rubber band to fire the plane into the air. The rubber band is stretched by 25 cm and has a spring constant of 120 N m^{-1}. The mass of the plane is 160 g. Assume that the rubber band follows Hooke's law and ignore its mass.

5 Calculate the magnitude of the force applied to the stretched rubber band to stretch it by 25 cm.

6 Calculate the strain potential energy stored in the stretched rubber band.

The following information relates to questions 7–10.
An Australian archer purchased a new bow for the Rio Olympics. Image (c) in the figure below shows the force required to pull back the 26 g arrow prior to launch. When no force is applied XY is 10 cm, as shown in image (a). Answer the following questions assuming the archer draws back the bow so that XY is 30 cm.



(a) **(b)** **(c)**

7 Calculate the strain potential energy stored in the stretched bow.

8 Calculate the work done by the archer.

9 Does the bow obey Hooke's law? Justify your answer.

10 Where on the graph is the elastic limit of the bow?

7.4 Kinetic and potential energy

FIGURE 7.4.1 The bungee jumper is in free fall until the cord starts to take up some of the kinetic energy and convert it to potential energy. The jumper is ultimately lowered safely to the ground.

A bungee jumper stakes their life on the physics principle of conservation of energy. The gravitational potential energy they lose as they jump off the bridge is transformed rapidly to kinetic energy, soon to be stored as strain potential energy in the bungee cord as the bungee jumper approaches the ground, and jerks back upwards relishing the adrenalin rush (see Figure 7.4.1). The calculations that predict that they will live to make another jump are the subject of this section.

KINETIC ENERGY

Kinetic energy (E_k) is the energy of motion of a body as you will recall from Chapter 5. For low speeds, it is calculated using:

$$E_k = \frac{1}{2}mv^2$$

In perfectly **elastic collisions**, kinetic energy is transferred between objects, and no energy is transformed into heat or sound or deformation. In cases such as these, the relationship is stated as:

$$E_k \text{ (before)} = E_k \text{ (after)}$$

In Chapter 5 you saw how momentum is always conserved in a collision. The total energy is also always conserved in a closed system; however in general, *kinetic* energy is not conserved, and such collisions are called **inelastic collisions**.

Perfectly elastic collisions do not exist in everyday situations, but they do exist in the interactions between atoms and subatomic particles. A collision between two billiard balls or the spheres in a Newton's cradle is almost perfectly elastic because very little of their kinetic energy is transformed into heat and sound energy.

Collisions such as a bouncing basketball, a gymnast on a trampoline and a tennis ball being hit are moderately elastic, with about half the kinetic energy of the system being retained. Perfectly inelastic collisions are those in which the colliding bodies stick together after impact with no kinetic energy. Some car crashes, a collision between a meteorite and the Moon, and a collision involving two balls of plasticine, could all be perfectly inelastic. In these collisions, most—and sometimes all—of the initial kinetic energy of the system is transformed into other forms of energy.

Worked example 7.4.1

ELASTIC OR INELASTIC COLLISION?

A car of mass 1.0×10^3 kg travelling west at 20 m s^{-1} crashes into the rear of a stationary bus of mass 5.0×10^3 kg. The vehicles lock together on impact. Show calculations to test whether or not the collision is inelastic.

Thinking	Working
Use conservation of momentum to find the final velocity of the wreck.	p_i (car) + p_i (bus) = p_f (car and bus) $1000 \times 20 + 0 = (1000 + 5000)\, v_f$ $v_f = 3.33$ m s^{-1}
Calculate the initial kinetic energy, before the collision for the bus and the car.	$E_{ki} = \frac{1}{2}mv_{ci}^2$ $= \frac{1}{2} \times 1.0 \times 10^3 \times 20^2$ $E_{ki} = 2.0 \times 10^5$ J
Calculate the final kinetic energy of the joined vehicles.	$E_{kf} = \frac{1}{2}m_{cb}v_f^2$ $= \frac{1}{2} \times (1.0 \times 10^3 + 5.0 \times 10^3) \times 3.33^2$ $E_{kf} = 3.3 \times 10^4$ J

Compare the kinetic energy before and after the collision to determine whether or not the collision is elastic.	The kinetic energy after the collision is significantly less than the kinetic energy before. The missing energy has been transformed to heat, sound and deformation of the vehicles. Therefore, this collision is inelastic.

Worked example: Try yourself 7.4.1

ELASTIC OR INELASTIC COLLISION?

A 200 g snooker ball with initial velocity 9.0 m s^{-1} to the right collides with a stationary snooker ball of mass 100 g. After the collision, both balls are moving to the right and the 200 g ball has a speed of 3.0 m s^{-1}. Show calculations to test whether or not the collision is inelastic.

POTENTIAL ENERGY

According the equation $E_g = mg\Delta h$, the **gravitational potential energy** of an object, E_g, is directly proportional to the mass of the object, m, its height above a reference point, Δh, and the strength of the gravitational field, g. This equation is derived from the fact that in order to lift an object of mass m through a distance Δh, work would be done against the force of gravity. Close to the surface of the Earth, this force is simply $F = mg$ and the distance travelled is $s = \Delta h$, so the work done is $W = Fs$, which is equal to the potential energy gained.

PHYSICSFILE

Energy equations

All equations for energy originate from the definition of work. If a force F acts on a body of mass m, causing a horizontal displacement s, the work done is:

$W = Fs = mas$

Now $v^2 = u^2 + 2as$ can be rearranged as:

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

$$W = ma\left(\frac{v^2 - u^2}{2a}\right)$$

$$= \frac{1}{2}mv^2 - \frac{1}{2}mu^2 = \Delta E_k \text{ (from the definition of work)}$$

In general, $E_k = \frac{1}{2}mv^2$.

Similarly, when a body is lifted at a uniform rate, the lifting force is simply equal to the gravitational force, i.e. mg. If the mass m is lifted through a vertical displacement s, the work done on the body is:

$W = Fs = mgs$

This vertical displacement is the change in height, Δh, in the gravitational field.

Thus:

$W = mg\Delta h = \Delta E_g$

In general, $E_g = mg\Delta h$.

When the gravitational force acting on an object can no longer be assumed to be constant, the gravitational potential energy can be calculated using a graph, in the same way that you calculated the work done by a varying force in sections 7.2 and 7.3. If the force is plotted as a function of distance, a graph like the one shown in Figure 7.4.2 is obtained (see page 246).

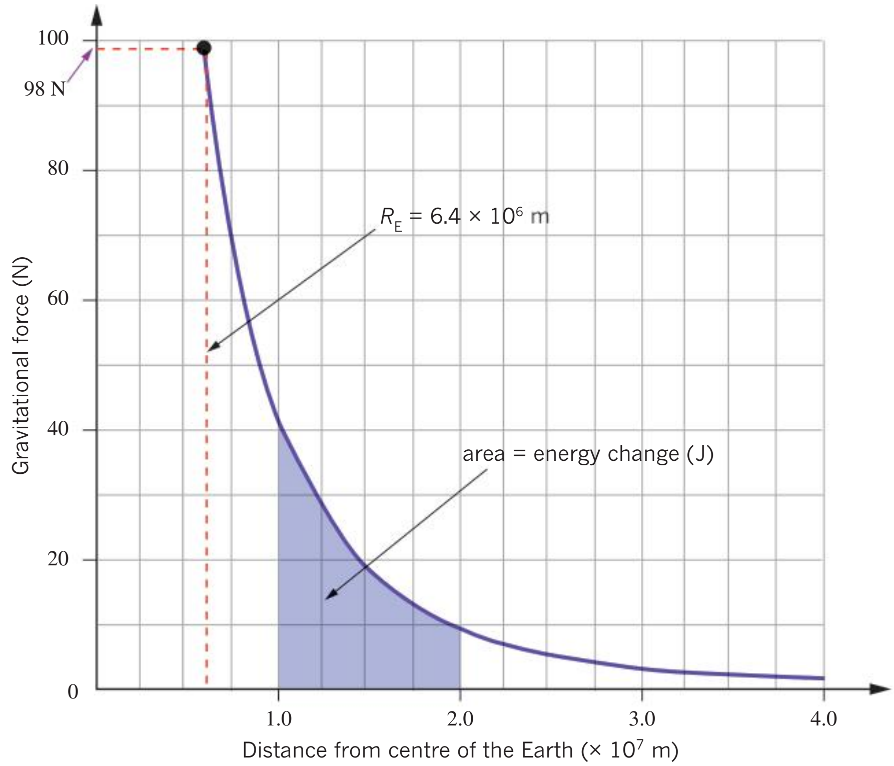


FIGURE 7.4.2 Plot of the gravitational force acting on a 10 kg body as a function of distance from the Earth. The shaded area represents the work done in moving the body a distance of 1.0×10^7 m.

Worked example 7.4.2

CHANGES IN GRAVITATIONAL POTENTIAL ENERGY USING A FORCE GRAPH

Using the graph in Figure 7.4.2, calculate the work done against the gravitational force in moving the 10 kg object from a radius of 1.0×10^7 m to 2.0×10^7 m, and hence find the gravitational potential energy gained.

Thinking	Working
Find the energy represented per square in the graph.	One square represents $10 \times 0.25 \times 10^7 = 2.5 \times 10^7$ J
Identify the two values of distance which are relevant to the question.	The object starts at 1.0×10^7 m and finishes at 2.0×10^7 m.
Count the squares under the curve between the two distance values identified above, and multiply by the energy per square.	Work done = potential energy gained 8.5 squares (approx) $\times 2.5 \times 10^7$ $= 2.1 \times 10^8$ J

Worked example: Try yourself 7.4.2

CHANGES IN GRAVITATIONAL POTENTIAL ENERGY USING A FORCE GRAPH

Use the graph in Figure 7.4.2 to estimate the gravitational potential energy of the 10 kg object relative to the surface of the Earth, for the 10 kg object at 2.0×10^7 m.

The disadvantage of the graph in Figure 7.4.2 is that it is specific to the mass of the object under consideration.

Recalling Newton's law of universal gravitation, the magnitude of the force between any two masses is given by

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

If the mass M is the object setting up the gravitational field (Earth in this example) then the graph above is found by calculating the force on the 10 kg object for each distance.

Sometimes it is more useful to create a graph of the force exerted per unit mass:

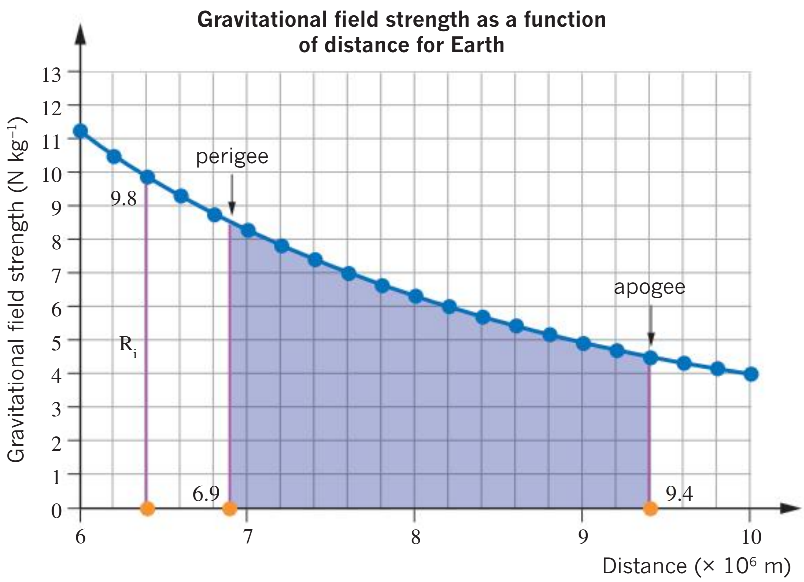
$$\frac{F_g}{m} = g = \frac{GM}{r^2}$$

This is often called the gravitational field strength, and it is dependent only on the body generating the field. Such a graph can be used to calculate the work done on any body in the field, as illustrated in Worked example 7.4.3.

Worked example 7.4.3

CHANGES IN GRAVITATIONAL POTENTIAL ENERGY USING A GRAVITATIONAL FIELD STRENGTH VS DISTANCE GRAPH

A wayward satellite of mass 1500 kg has developed a highly elliptical orbit around the Earth. At its closest approach (perigee), it is just 500 km above the Earth's surface. Its furthest point (apogee) is 3000 km from the Earth's surface. The Earth has a mass of 6.0×10^{24} kg and a radius of 6.4×10^6 m. The gravitational field strength of the Earth is shown in the graph.



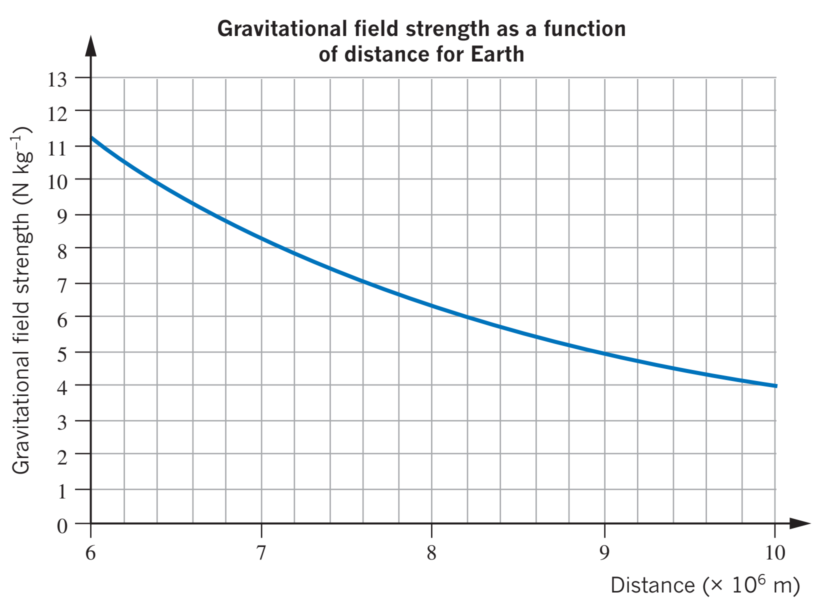
a Calculate the change in potential energy of the satellite as it moves from the closest to the furthest point from the Earth.	
Thinking	**Working**
Convert distances given as altitudes to distances from the centre of Earth.	Perigee = $6.4 \times 10^6 + 500 \times 10^3$ $= 6.9 \times 10^6$ m Apogee = $6.4 \times 10^6 + 3000 \times 10^3$ $= 9.4 \times 10^6$ m
Find the energy represented by each square.	One square represents $1.0 \times 0.20 \times 10^6 = 2.0 \times 10^5$ J kg^{-1}
Count the squares under the curve for the relevant area, and multiply by the energy per kg represented by each square.	Work done per kg = potential energy gained per kg of mass 76 squares (approx) $\times 2.0 \times 10^5$ $= 1.52 \times 10^7$ J kg^{-1}
Calculate the potential energy gained for the satellite by multiplying by the mass of the satellite.	Energy gained $E_g = 1.52 \times 10^7 \times 1500$ $= 2.3 \times 10^{10}$ J

b The satellite was moving with a speed of 15 km s^{-1} at its closest point to Earth. How fast was it travelling at its furthest point?	
Thinking	**Working**
First calculate the kinetic energy at the closest point (perigee).	$E_{kp} = \frac{1}{2}mv_p^2$ $= \frac{1}{2} \times 1500 \times (15 \times 10^3)^2$ $= 1.69 \times 10^{11}$ J
The gain in gravitational potential energy at the furthest point (apogee) is at the expense of kinetic energy. Calculate the kinetic energy at the apogee.	$E_{ka} = E_{kp} - E_g$ $= 1.69 \times 10^{11} - 2.3 \times 10^{10}$ $= 1.46 \times 10^{11}$ J
Calculate the speed of the satellite at the apogee.	$E_{ka} = \frac{1}{2}mv_a^2$ $v_a = \sqrt{\frac{2E_{ka}}{m}}$ $= \sqrt{\frac{2 \times 1.46 \times 10^{11}}{1500}}$ $= 14$ km s^{-1}

Worked example: Try yourself 7.4.3

CHANGES IN GRAVITATIONAL POTENTIAL ENERGY USING A GRAVITATIONAL FIELD STRENGTH VS DISTANCE GRAPH

A satellite of mass 1100 kg is in an elliptical orbit around Earth. At its closest approach (perigee), it is just 600 km above Earth's surface. Its furthest point (apogee) is 2600 km from Earth's surface. The Earth has a mass of 6.0×10^{24} kg and a radius of 6.4×10^6 m. The gravitational field strength of Earth is shown in the graph.



a Calculate the change in potential energy of the satellite as it moves from the closest to the furthest point from Earth.

b The satellite was moving with a speed of 8.0 km s^{-1} at its closest point to Earth. How fast was it travelling at its furthest point?

WORK AND ENERGY

Both work and energy are scalar quantities, and as such have only magnitude. It is important, however, that you keep account of whether energy is being gained or lost by an object. If work is being done by a body, it loses kinetic or potential energy. If work is done on the body by an external force, the body would gain energy.

A weightlifter loses chemical potential energy as he exerts a force on a barbell to lift the bar. If he lifts the bar at constant speed, the bar does not gain kinetic energy, but gains gravitational potential energy. In drawing back an arrow, an archer does work on the bow and string, and this strain potential energy is transformed to the kinetic energy of the arrow when the string does work on the arrow as it is released (see Figure 7.4.3).

FIGURE 7.4.3 The archer does work on the bow, and strain potential energy is stored. This is later transformed into the kinetic energy of the arrow.

For a change in gravitational potential energy, the reference point from which the change in height Δh is measured is arbitrary. A reference position on the surface of the Earth is frequently used for gravitational potential energy close to the surface of the Earth. It is possible, in this sense, to get a 'negative potential energy'. For instance, if you were to take the surface of a table as the zero for potential energy, raising an object above the surface would require work to be done and it would increase in energy. Positioning the object on the floor below the table would amount to its having less potential energy than on the table, effectively giving it a negative potential energy. Therefore it is only changes in gravitational potential energy that are meaningful for your purposes, and the sign merely indicates whether energy has been gained or lost relative to the zero point.

Worked example 7.4.3 implicitly assumed zero gravitational potential energy to be on the surface of Earth. Putting a satellite into circular orbit requires doing work to increase its gravitational potential energy, because a force has to be applied to oppose gravity. The higher the orbit, the greater the gravitational potential energy relative to the surface. This was calculated in Worked example 7.4.3 from the area under the gravitational force vs distance curve.

The kinetic energy of a given satellite in a circular orbit decreases with increasing altitude. This is shown by considering the gravitational force on a satellite of mass m from a central body of mass M to be equal to the centripetal force:

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

From which can be found that $v = \sqrt{\frac{GM}{r}}$ and $E_k = \frac{1}{2}\frac{GMm}{r}$.

ENERGY TRANSFORMATIONS

The sum of the potential and kinetic energy of an object is its **mechanical energy**, and this is constant unless work is done by an external force. There is frequently transformation of energy between potential and kinetic energy. A child dropping from the branch of a tree onto a trampoline loses gravitational potential energy, but gains kinetic energy. On striking the trampoline, kinetic energy is transformed to strain potential energy in the springs, and in an ideal case, would be returned as kinetic energy on the rebound.

Sometimes energy is dissipated, or transformed into heat, light and/or sound, and thus the energy remaining in the system is reduced. Spacecraft have to dissipate huge amounts of kinetic and gravitational potential energy as they re-enter the Earth's atmosphere and slow down to make a landing. Meteors, or so-called shooting stars, burn up in the upper atmosphere because of the heat generated by friction.

7.4 Review

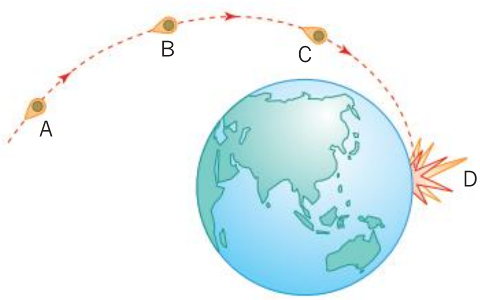
SUMMARY

- Kinetic energy is the energy of motion of a body: $E_k = \frac{1}{2}mv^2$
- The sum of the kinetic and potential energy (total mechanical energy) of an isolated system is always conserved.
- For perfectly elastic collisions, the kinetic energy before the collision is equal to the kinetic energy after the collision.
- Close to the surface of the Earth, where the force of gravity can be assumed to be constant, the change in gravitational potential energy for an object of mass m is $E_g = mg\Delta h$, where the height changes by Δh.
- For a non-constant gravitational force, the gravitational potential energy can be calculated from the area under a graph of force vs distance.
- For convenience, force–distance graphs are often plotted as force per unit mass vs distance, so that the same graph can be used for any mass. In this case the area is energy per unit mass.
- When work is done on a body it gains mechanical energy.
- When the body does work, energy is dissipated to the environment, for example as heat, sound or deformation, and the body loses mechanical energy.

KEY QUESTIONS

1 Calculate the gravitational potential energy of a 115 kg climber standing at the top of Mount Kosciuszko when he is at an altitude of 2228 m above sea level.

2 The figure below shows a meteor plunging towards Earth, partially burning up in the atmosphere as it does so.
Choose which statements are correct. More than one correct answer is possible.



A The kinetic energy of the meteor increases as it travels from A to D.

B The gravitational potential energy of the meteor relative to the surface of the Earth increases as it travels from A to D.

C The total energy of the meteor increases as it travels from A to D.

D The total mechanical energy of the meteor remains constant.

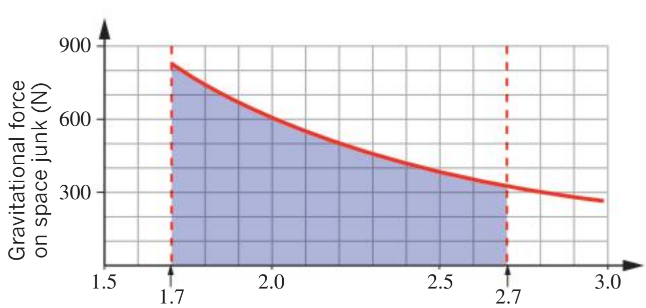
E The gravitational potential energy of the meteor relative to the surface of the Earth decreases as it travels from A to D.

The following information relates to questions 3–5.

A 500 kg lump of space junk is plummeting towards the Moon. Its speed when it is 2.7×10^6 m from the centre of the Moon is 250 m s^{-1}.

The Moon has a radius of 1.7×10^6 m.

The gravitational force–distance graph is shown below:

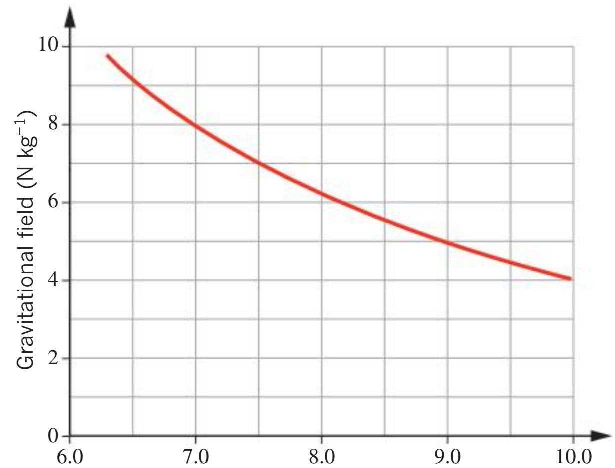


3 Calculate the kinetic energy of the junk when it is travelling at 250 m s^{-1}.

4 Calculate the increase in kinetic energy of the junk as it falls to the Moon's surface.

5 Calculate the speed of the junk as it crashes into the Moon.

6 A 20 tonne satellite is in orbit at an altitude of 600 km. A booster rocket is fired putting the satellite into an orbit at an altitude of 2600 km. Calculate the work done by the booster rocket to increase the potential energy of the satellite using the graph below.
Assume the radius of Earth is 6.4×10^6 m.



7 A cable car system is set up so that the cars of the same mass are attached to a moving cable that is powered by a motor at one end. As car A is pulled upwards, car B descends, both at the same speed. Select the statements that are correct. More than one correct answer is possible.

A Car A and car B each have constant kinetic energy.

B Car A and car B each have constant gravitational potential energy.

C As the gravitational potential energy of car A increases, that of car B decreases.

D The motor does work on the cable.

8 The 11-tonne Hubble telescope is in a circular orbit at an altitude of approximately 600 km above the surface of Earth. A geosynchronous weather satellite of the same mass is in an orbit at an altitude of approximately 3600 km. Select the statements that are correct. More than one correct answer is possible.

A The gravitational potential energy of the geosynchronous satellite is 6 times that of the Hubble telescope, relative to the surface of the Earth.

B The Hubble telescope's orbital speed is greater than that of the weather satellite.

C The kinetic energy of the weather satellite is greater than that of the Hubble telescope.

D The weather satellite has more gravitational potential energy than the Hubble telescope, relative to the surface of Earth.

9 If a high-jumper with a mass of 63 kg just clears a height of 2.1 m, what was the high jumper's speed as he left the ground?

7.5 Einstein's mass–energy relationship

FIGURE 7.5.1 Mass is relative to the frame of reference in which it is measured.

PHYSICSFILE

Travel at the speed of light

Einstein said that at the speed of light distances shrink to zero and time stops. No ordinary matter can reach c, but light always travels at c. Strange though it may seem, for light there is no time. It appears in one place and disappears in another, having got there in no time (in its own frame of reference, not ours!). When you stay still, you travel through spacetime in the time dimension only. Light does the opposite: all its spacetime travel is through space and none through time.

This chapter refers to both energy and momentum being conserved. In classical physics and chemistry the conservation of mass is assumed: the particles that start out in a chemical reaction are still there at the end, and applying a force to an object classically does not change its mass. Now it is time to look at the implications of Einstein's relativistic principles, and to show the development of Einstein's most famous equation relating energy, mass and the speed of light, $E = mc^2$. In order to do this, this section will first look at what happens to the momentum of an object as its speed approaches the speed of light.

APPROACHING THE SPEED OF LIGHT

Recall the Lorentz factor that was introduced in Chapter 6:

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

At low speeds, γ is so close to 1 that the effects of special relativity can be ignored, but γ rapidly increases as the speed, v, comes closer to the speed of light, c. At 99.9% of the speed of light, γ has a value of approximately 22, and so anything moving at that speed, relative to a stationary observer, will appear to have shrunk to $\frac{1}{22}$ of its normal length. As you watch the action inside a spaceship travelling at that speed, events would appear to be going 22 times more slowly than they would if they occurred in a stationary observer's frame of reference.

The closer that the speed of the spaceship gets to the speed of light, the more the Lorentz factor increases towards infinity. It is reasonable to wonder what happens at the speed of light. According to Einstein's equations, the length of the spaceship shrinks to zero and time inside it appears to stop altogether. Einstein took this to mean that it is not possible to reach the speed of light in any real spaceship. However, the difficulties with time and length for the spaceship were not the only reasons Einstein came to this conclusion.

Relativistic momentum

If a rocket ship like the one in Figure 7.5.2 is travelling at $0.99c$, why can't it simply turn on its rocket motor and accelerate up to c, or more? A full answer to this question was not given in Einstein's original 1905 paper on relativity. Some years later he showed that as the speed of a spaceship approaches c, its momentum increases, but this is not reflected in a corresponding increase in speed.

Although his analysis is beyond the scope of this course, you can get a feel for his approach if you take some short cuts.

The acceleration a, of any object is inversely proportional to its mass m, the mass that appears in Newton's second law:

$$F = ma$$

Newton originally stated this law as: a force, F, is equal to the rate of change in momentum p. That is:

$$F = \frac{\Delta p}{\Delta t}$$

A change in momentum is classically defined as the change in the product of the mass, m, and the velocity, v. If you rearrange the above equation and substitute the relationship $\Delta p = m\Delta v$, you get:

$$F\Delta t = m\Delta v$$

Now you see that time is involved, but at relativistic speeds you know that time is not the constant entity it was once believed to be.

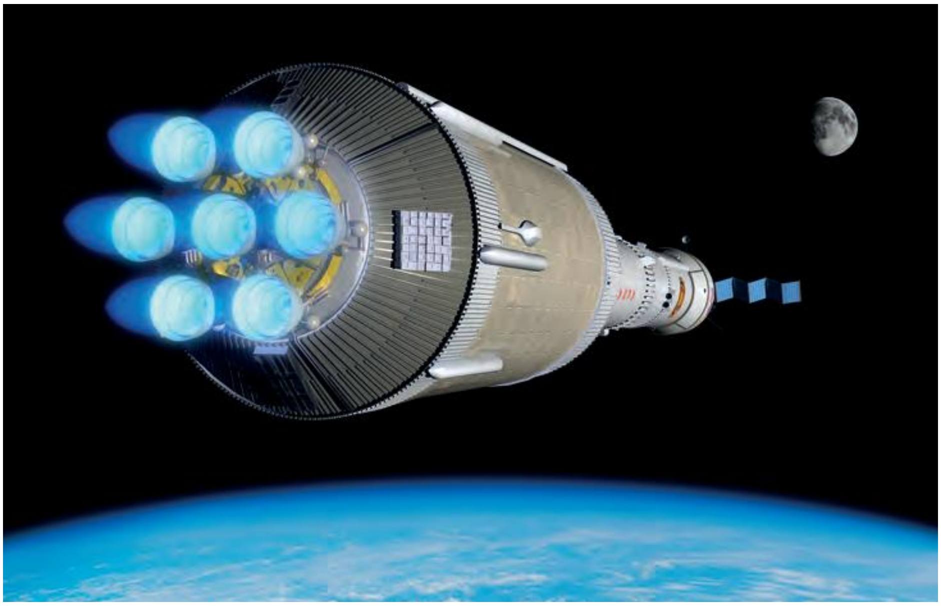


FIGURE 7.5.2 This rocket ship is moving at 0.99*c* and accelerating, and yet it can never reach a speed of *c*.

Imagine that you have a rocket ship accelerating from rest to a high speed as viewed by an observer in a stationary frame of reference. You can say that the change in momentum of the ship will be given by:

$F\Delta t_0 = m\Delta v$

where t_0 is the time in the ship's frame of reference, and $m\Delta v$ is just the classical Newtonian change in momentum.

In the stationary observer's frame, the time is dilated:

$\Delta t = \gamma \Delta t_0$

$\Delta t_0 = \frac{\Delta t}{\gamma}$

Substituting Δt_0 into the change of momentum equation above:

$F\frac{\Delta t}{\gamma} = m\Delta v$

$F\Delta t = \gamma m\Delta v$

That is, the impulse as seen by the stationary observer is equal to the product of the Lorentz factor, γ, and the change in Newtonian momentum. This means that as the spaceship approaches the speed of light, the impulse is multiplied by a factor that grows very rapidly. You can interpret this as meaning that the change in momentum in the stationary observer's frame of reference is equal to:

$\Delta p = \gamma m\Delta v$

$\Delta p = \gamma \Delta p_0$

If we assume an object starts at zero velocity, the final relativistic momentum becomes:

$p = \gamma mv$

$p = \gamma p_0$

where p_0 is the momentum mv, as you would define it in classical mechanics, and p is the relativistic momentum.

If velocity, v, is needed when the mass and relativistic increase in momentum is known, the formula $p = \gamma mv$ can be rearranged to give the following:

$$v = \frac{p}{m\sqrt{\left(1 + \frac{p^2}{m^2c^2}\right)}}$$

EXTENSION

Rearranging relativistic momentum to determine velocity

$p = \gamma mv$

$$p = \frac{1}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}} mv$$

$$\sqrt{\left(1 - \frac{v^2}{c^2}\right)} = \frac{mv}{p}$$

$$1 - \frac{v^2}{c^2} = \left(\frac{mv}{p}\right)^2$$

$$1 = \left(\frac{mv}{p}\right)^2 + \frac{v^2}{c^2}$$

$$\left(\frac{m^2}{p^2} + \frac{1}{c^2}\right)v^2 = 1$$

$$\frac{m^2}{p^2}\left(1 + \frac{p^2}{m^2c^2}\right)v^2 = 1$$

$$v^2 = \frac{p^2}{m^2\sqrt{\left(1 + \frac{p^2}{m^2c^2}\right)}}$$

$$v = \frac{p}{m\sqrt{\left(1 + \frac{p^2}{m^2c^2}\right)}}$$

While there are some short cuts taken to reach this result, the result itself is perfectly valid.

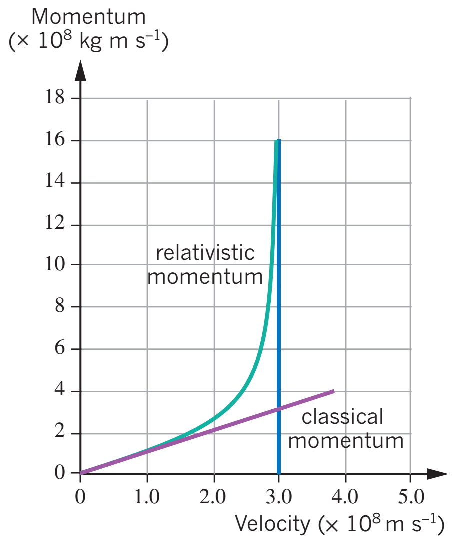


FIGURE 7.5.3 The relationship between classical momentum and velocity, and the relationship between relativistic momentum and velocity, for a 1 kg mass.

The momentum increases very rapidly as a spaceship approaches the speed of light. You might argue that this is expected—after all, momentum is a function of velocity. If you graph the relativistic momentum, p, against the velocity, v, and on the same graph show the classical momentum, you can see that the relativistic momentum increases at a rate far greater than it would if it were due to the increase in velocity alone (see Figure 7.5.3).

This result can be interpreted by thinking of the mass as a quantity that also increases at high speeds. Thus there is a relationship between the rest mass, m, which is the mass measured while the object is at rest in the frame of reference, and the relativistic mass, γm, which is the mass measured as the object is moving relative to the observer.

As the Lorentz factor increases with the increase in the velocity, then the relativistic mass also increases. Einstein was never happy with the term 'relativistic mass', and preferred that people only spoke of the relativistic momentum of an object.

Notice too how the classical treatment allows the object to have a speed greater than the speed of light, but the relativistic treatment causes the mass to become very large so that the speed of light is never actually reached.

Now return to the example of the rocket ship that is attempting to increase its velocity to the speed of light. With the increase in the relativistic mass of the rocket ship, it becomes harder for the force of the engines to cause a change in velocity. The closer the rocket ship approaches c, the greater the amount of impulse that is required to accelerate the ship to the speed of light. In fact, as the velocity approaches c, the relativistic mass, γm, approaches infinity. You can now see why your rocket ship cannot reach the speed of light.

Worked example 7.5.1 illustrates this point. Notice that the result in part (b) shows that if you double the impulse required to get the rocket ship to 0.99c, then you will only add 0.007c to your top speed. When you've completed Worked example: Try yourself 7.5.1, consider the change in velocity achieved by tripling the impulse.

Worked example 7.5.1

RELATIVISTIC MOMENTUM

a Calculate the momentum, as seen by a stationary observer, provided to a rocket ship with a rest mass of 1000 kg, as it goes from rest up to a speed of 0.990c. Assume *Gedanken* conditions exist in this example.

Thinking	Working
Identify the variables: the rest mass is m, and the velocity of the rocket ship is v.	$\Delta p = ?$ $m = 1000$ kg $v = 0.990 \times 3.00 \times 10^8$
Use the relativistic momentum formula.	$p = \gamma mv$
Substitute the values for m and v into the equation and calculate the answer p.	$p = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} mv$ $= \frac{1}{\sqrt{1 - \frac{0.990^2 c^2}{c^2}}} \times 1000 \times 0.990 \times 3.00 \times 10^8$ $= 2.11 \times 10^{12}$ kg m s^{-1}

PHYSICSFILE

Getting beyond *c*

As you watch a rocket ship travelling at 0.99c (speed U), it fires a small ship at 0.02c relative to it (speed v). Isn't the small ship moving at 1.01c? No! First you need to be careful to specify in which frame of reference you are measuring the speeds. The rocket ship has speed U in your frame while the small ship has speed v in the frame of the rocket ship. (Capital letters for your frame, small for the rocket frame.) Because of length contraction you see the small ship fired at much less than 0.02c. Einstein showed that in these cases the speed (V) of the small ship in your frame is given by:

$$V = \frac{U + v}{1 + \frac{Uv}{c^2}}$$

You can use this expression to show that we see the small ship travelling at 0.9904c.

b If twice the relativistic momentum from part a is applied to the stationary rocket ship, calculate the new final speed of the rocket ship in terms of *c*.	
Thinking	**Working**
Identify the variables: the rest mass is *m*, and the relativistic momentum of the rocket ship is *p*.	$p = 2 \times 2.11 \times 10^{12}$ $= 4.21 \times 10^{12}$ kg m s^{-1} $m = 1000$ kg $v = ?$
Use the relativistic momentum formula, rearranged.	$p = \gamma mv$ $p = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} mv$ $v = \frac{p}{m\sqrt{\left(1 + \frac{p^2}{m^2c^2}\right)}}$
Substitute the values for *m* and *p* into the rearranged equation and calculate the answer *v*.	$v = \frac{p}{m\sqrt{\left(1 + \frac{p^2}{m^2c^2}\right)}}$ $= \frac{4.21 \times 10^{12}}{(1000)\sqrt{1 + \frac{(4.21 \times 10^{12})^2}{1000^2(3.00 \times 10^8)^2}}}$ $= \frac{4.21 \times 10^{12}}{1000 \times 14.07}$ $= 2.99 \times 10^8$ m s^{-1} $= 0.997c$

Worked example: Try yourself 7.5.1

RELATIVISTIC MOMENTUM

a Calculate the momentum, as seen by a stationary observer, provided to an electron with a rest mass of 9.11×10^{-31} kg, as it goes from rest to a speed of 0.985*c*. Assume *Gedanken* conditions exist in this example.

b If three times the relativistic momentum from part (a) is applied to the electron, calculate the new final speed of the electron in terms of *c*.

EINSTEIN'S FAMOUS EQUATION

As the momentum of an object increases, so does its kinetic energy. The classical relationship between the two can be written as:

$$E_k = \frac{1}{2}mv^2$$
$$= \frac{1}{2}(mv) \times v$$
$$= \frac{1}{2}pv$$

This form of the equation shows that the kinetic energy of an object is related to an object's momentum as well as its velocity.

Einstein showed, however, that the classical expression for kinetic energy was not correct at high speeds. The mathematics involved is beyond the scope of this course, but Einstein, working from the expression for relativistic momentum and the usual assumptions about work, forces and energy, was able to show that the kinetic energy of an object was given by the expression:

$$E_k = (\gamma - 1)mc^2$$

Although it is not very obvious from this expression, if the velocity (which is hidden in the γ term) is small, this expression actually reduces to the classical equation for E_k of $\frac{1}{2}mv^2$. A small velocity in this context means small in comparison to *c*. But even for speeds up to 0.10*c*, the classical expression is accurate to better than ±1% .

Einstein's expression can be expanded to:

$E_k = \gamma mc^2 - mc^2$

This kinetic energy equation, in turn, can be rearranged as:

$\gamma mc^2 = E_k + mc^2$

Einstein interpreted the left-hand side of this expression as being an expression for the total energy of the object:

$E_{tot} = \gamma mc^2$

The right-hand side appeared to imply that there were two parts to the total energy: the kinetic energy, E_k, and another term that only depended on the rest mass, m. The second term, mc^2, he referred to as the rest energy of the object, as it does not depend on the speed of the object. This appeared to imply that somehow there was energy associated with mass (see Figure 7.5.4). An astounding proposition to a classical physicist, but as you have seen, in relativity, mass increases as you add kinetic energy to an object. The conservation of energy relationship is therefore:

$E_{tot} = E_k + E_{rest}$

$E_{tot} = \gamma mc^2$

where γm is the relativistic mass (kg)

c is the speed of light (m s^{-1}) and

E_{total} is the total energy (J)

You will have seen part of this equation before:

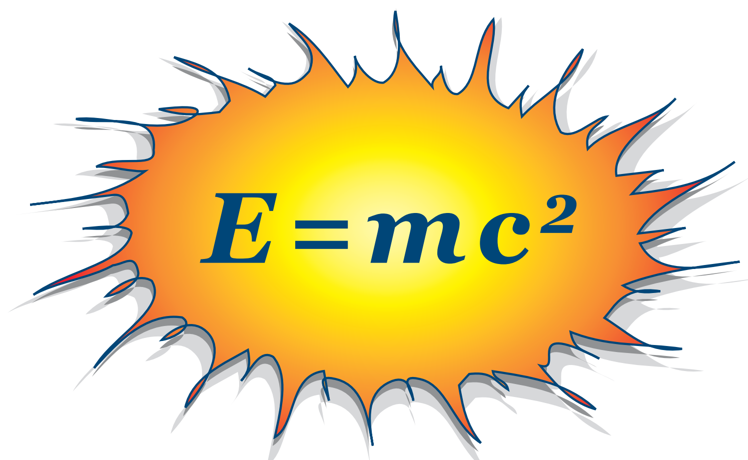


FIGURE 7.5.4 Einstein's famous equation.

This equation tells you that mass and energy are totally interrelated. In a sense, you can say that mass has energy, and energy has mass.

CONVERTING MASS TO ENERGY OR ENERGY TO MASS

Nuclear reactions involve vastly more energy per atom than chemical ones (see Figure 7.5.5). When a uranium atom splits into two fission fragments, about 200 million electron volts of energy are released. By comparison, most chemical reactions involve just a few electron volts.

In the fission of uranium, it is possible to find the original mass of the uranium nucleus and the fission fragments accurately enough to determine the mass defect (change in mass). This difference in mass agrees exactly with the prediction of Einstein's famous equation.

Likewise, nuclear fusion reactions deep inside the Sun release the huge amounts of energy that stream from the Sun, resulting in a conversion of about 4 million tonnes of mass into energy every second (see Figure 7.5.6).

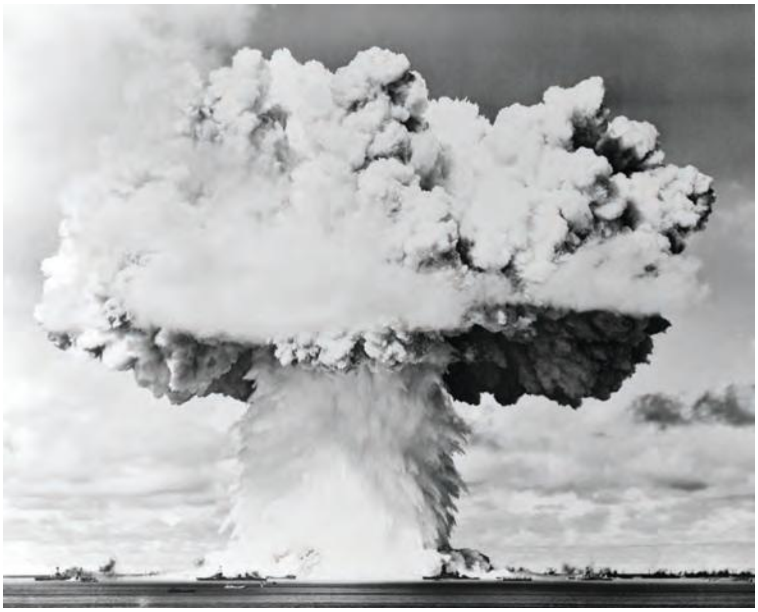


FIGURE 7.5.5 In a nuclear bomb, a few grams of mass are converted into energy. As the uranium undergoes fission, it releases the equivalent of hundreds of gigajoules (10^{12} J) of energy. Millions of tonnes of a chemical explosive (TNT) would be required to produce this much explosive energy.

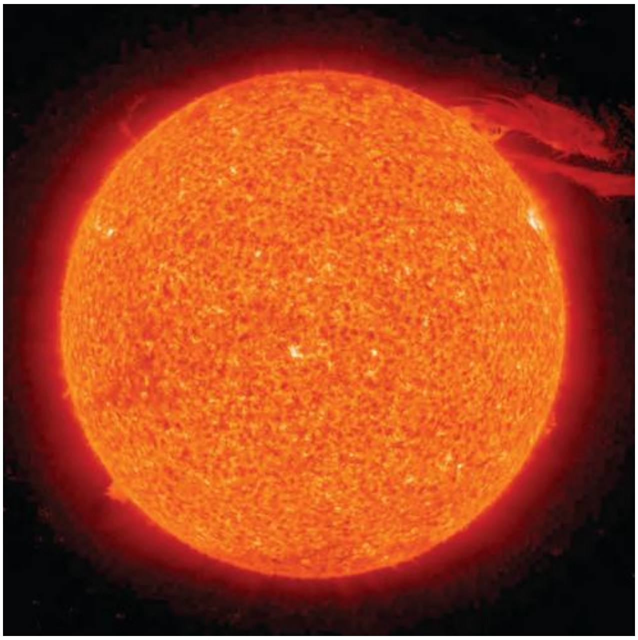


FIGURE 7.5.6 Nuclear fusion in the Sun results in about 4 million tonnes of mass being converted into energy every second, which is radiated from the Sun.

NUCLEAR FUSION

Nuclear fusion occurs when two light nuclei are combined to form a larger nucleus (see Figure 7.5.7).

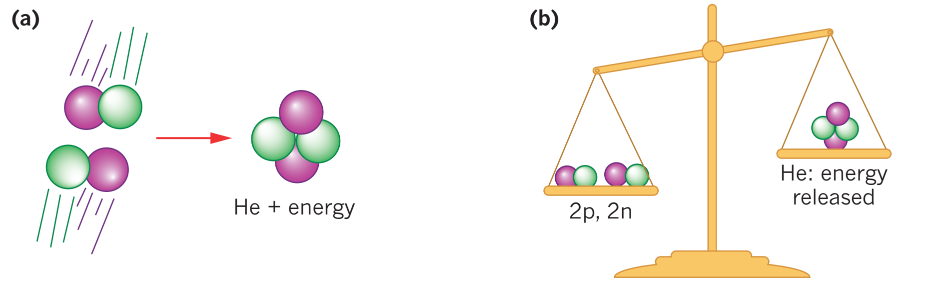


FIGURE 7.5.7 (a) When two isotopes of hydrogen fuse to form a helium nucleus, energy is released. (b) The binding energy of the nucleus appears as a loss in mass, Δm, which can be calculated using $\Delta E = \Delta mc^2$.

As in the cases of radioactive decay and nuclear fission, the mass of the reactants is slightly greater than the mass of the products when the nuclei combine during fusion.

The energy created by this missing mass can again be determined from:

$$\Delta E = \Delta mc^2$$

where ΔE is the energy (J)

Δm is the mass defect (kg)

c is the speed of light (3.0×10^8 m s^{-1}).

Nuclear fusion is a very difficult process to recreate in a laboratory. The main problem is that nuclei are positively charged, and thus repel one another.

Slow-moving nuclei with relatively small amounts of kinetic energy will not be able to get close enough for the strong nuclear force to come into effect, and so fusion will not happen. Only if nuclei have enough kinetic energy to overcome the repulsive force can they come close enough for the strong nuclear force to start acting. If this happens, fusion will occur (see Figure 7.5.8).

Typically, temperatures of the order of hundreds of millions of degrees are required. These are exactly the conditions that are present inside the Sun.

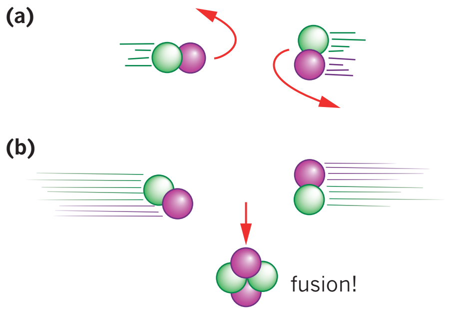
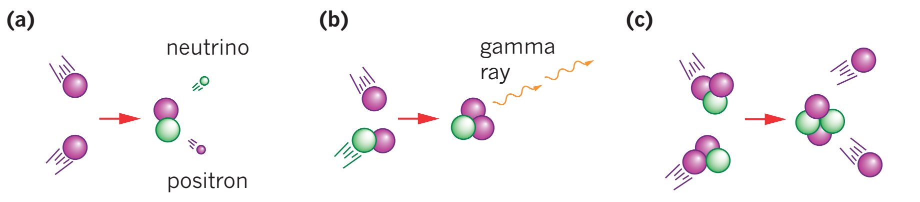


FIGURE 7.5.8 (a) Slow-moving nuclei do not have enough energy to fuse together. The electrostatic forces cause them to be repelled from each other. (b) If the nuclei have sufficient kinetic energy, then they will overcome the repulsive forces and move close enough together for the strong nuclear force to come into effect. At this point, fusion will occur and energy will be released.

FUSION IN THE SUN AND SIMILAR STARS

In the Sun, many different fusion reactions are taking place. The main reaction is the fusion of hydrogen nuclei to form helium. Each second, about 657 million tonnes of hydrogen and hydrogen isotopes fuse to form about 653 million tonnes of helium. Each second, a mass defect of 4 million tonnes results from these fusion reactions. The amount of energy released is enormous and can be found by using the equation $\Delta E = \Delta mc^2$. A tiny proportion of this energy reaches Earth and sustains life as we know it.

The sequence of fusion reactions shown in Figure 7.5.9 has been occurring inside the Sun for the past 5 billion years and is expected to last for another 5 billion years or so. Hydrogen nuclei are fused together and, after several steps, a helium nucleus is formed. This process releases about 25 MeV of energy.



(a) ${}^{1}_{1}H + {}^{1}_{1}H \longrightarrow {}^{2}_{1}H + {}^{0}_{+1}e + {}^{0}_{0}\nu$

(b) ${}^{1}_{1}H + {}^{2}_{1}H \longrightarrow {}^{3}_{2}He + {}^{0}_{0}\gamma$

(c) ${}^{3}_{2}He + {}^{3}_{2}He \longrightarrow {}^{4}_{2}He + {}^{1}_{1}H + {}^{1}_{1}H$

FIGURE 7.5.9 The three main fusion reactions taking place inside the Sun

The Sun is a second- or third-generation star. It was formed from the remnants of other stars that exploded much earlier in the history of the galaxy. As this giant gas cloud contracted under the effect of its own gravity, the pressure and temperature at the core reached extreme values, sufficient to sustain these fusion reactions.

Worked example 7.5.2

FUSION

Consider the fusion reaction shown below. A proton fuses with a deuterium nucleus (a hydrogen nucleus with one neutron) in the Sun. A helium nuclide is formed and a γ-ray released. 20 MeV of energy is released during this process.

$${}^{1}_{1}p + {}^{2}_{1}H \rightarrow {}^{x}_{2}He + \gamma$$

a What is the value of the unknown mass number x?

Thinking	Working
Analyse the mass numbers. The gamma ray has atomic and mass numbers of zero.	$1 + 2 = x + 0$ $x = 3$ A helium-3 nucleus is formed.

b How much energy is released in joules?

Thinking	Working
$1\ \text{eV} = 1.6 \times 10^{-19}\ \text{J}$	$20\ \text{MeV} = 20 \times 10^{6} \times 1.6 \times 10^{-19}$ $= 3.2 \times 10^{-12}\ \text{J}$

c Calculate the mass defect for this reaction.

Thinking	Working
Use $\Delta E = \Delta mc^2$.	$\Delta m = \frac{\Delta E}{c^2}$ $= \frac{3.2 \times 10^{-12}}{(3.0 \times 10^{8})^2}$ $= 3.6 \times 10^{-29}\ \text{kg}$

Worked example: Try yourself 7.5.2

FUSION

A further fusion reaction in the Sun fuses two helium nuclides. A helium nucleus and two protons are formed and 30 MeV of energy is released.

$^{3}_{2}\text{He} + ^{3}_{2}\text{He} \rightarrow ^{x}_{2}\text{He} + 2^{1}_{1}\text{H}$

a What is the value of the unknown mass number *x*?
b How much energy is released in joules?
c Calculate the mass defect for this reaction.

PHYSICS IN ACTION

Magnetism and relativity

Magnetism is a relativistic phenomenon! You may think that relativity only applies to ultrafast rocket ships or exotic particles in accelerators, but, in fact, before relativity there was no good explanation of the magnetic forces between electric currents moving at speeds of only millimetres per second. The 19th-century physicists knew that there was a problem with the theory of magnetism. It was well known that a moving charge (and hence a current) in a magnetic field experiences a force that is directly proportional to the velocity of the charge ($F = BIl = qvB$, see Chapter 3, page 96). Indeed, this is the force that drives all the electric motors of the world. But how could there be a velocity-dependent force without contravening the Galilean principle of relativity?

The problem was that in the frame of reference of the moving charge, the velocity of the charge, and hence the magnetic force on it, should be zero. This was clearly not the case, however. Physicists pondered whether there was something special about the frame of reference after all, despite the Galilean principle of relativity. At the beginning of the 20th century this was one of the unsolved mysteries of physics.

Consider two electric currents moving in the same direction in two similar parallel wires. In this situation there is a magnetic force of attraction between the two wires; a simple experiment can confirm this. Now, imagine a moving electron in one of the wires. It 'sees' the magnetic field created by the current in the other wire and 'feels' a magnetic force towards it. If this situation was observed from a frame of reference moving at the same velocity as the electrons, (which is literally only a snail's pace) the electrons would all be at rest and so there should be no force!

Now, the force between two objects cannot depend on the frame of reference. Either they will get closer or they won't, and that doesn't depend on how you look at them. Wires with parallel currents do get closer, so there would be a fault in the physics if it said there is no force between them. This type of situation is not uncommon in science. In fact, it is one of the ways in which science progresses. Einstein was very aware of the problem of electromagnetism and indeed his famous 1905 paper starts with a discussion of just this problem.

Normally, when discussing electric currents you simply think of the moving electrons and ignore the huge numbers of positive and negative charges that are at rest in the wire. The very good reason for this assumption is that the total negative charge of the conduction electrons is almost equal in magnitude to the total positive charge of the positive ions (the atoms with the remaining electrons); so the wire as a whole is neutral. (Any small overall charge because of a positive or negative voltage on the wire is negligible.) As in any problem, however, you need to look at your assumptions very carefully.

Our hypothetical moving electron actually 'sees' a huge electrostatic (Coulomb) force towards all the positive charges in the other wire, but this is balanced by the equally huge repulsive force from all the negative charges in that wire—or is it? Classical theory certainly says that these forces should balance. However, relativity tells the physicist to be careful where there is relative motion. The moving electron actually sees all the moving electrons in the other wire at rest, relative to itself, but it sees all the positive charges moving in the opposite direction. Now that means that the positive charges will appear—to our moving electron—contracted or, more particularly, the space that they occupy will appear to be shortened in the direction of their motion.

And so their density—the number of positive charges in a metre of wire—appears greater than the density of the negative charges (see Figure 7.5.10(c)). There is, therefore, an imbalance in the Coulomb force between our electron and the negative and positive charges in the other wire. The electron sees more positive charges than negative charges and so is attracted to the other wire!

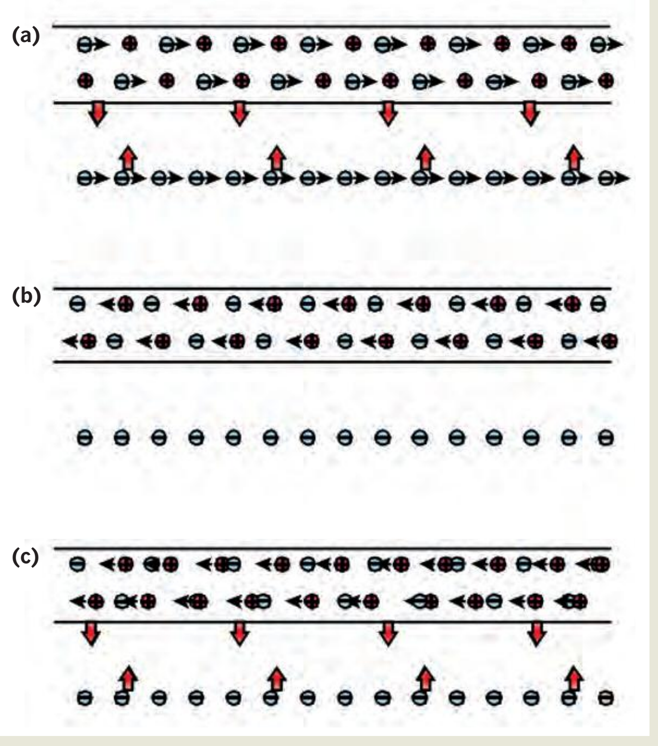


FIGURE 7.5.10 These three diagrams represent two similar electric currents. (a) This is the conventional view; the stream of electrons and the current are attracted because the moving electrons in each create a magnetic field and experience a force from the field created by the other. (b) This is the view from the frame of reference of the electrons. (c) This is the relativistic version.

The obvious comment to be made here, however, is that the speed involved—a few millimetres per second—is so small that the Lorentz contraction should be totally negligible. However, it is important to remember that you are dealing with huge numbers of charges. It is worth doing a simple calculation as an illustration of the forces you are dealing with:

If you could take all the conduction electrons out of a piece of wire and separate them by 1 cm from all the positive ions remaining, what would be the force between them? A piece of copper wire 10 cm long and with a cross-sectional area of 1 mm^2 has a mass of about 1 g. As its atomic mass is 64 atomic mass units, there are about 10^{22} atoms. Assume that one electron from each atom is taken out and placed 1 cm away from the remaining positive ions. The total charge of the electrons is then $10^{22} \times 1.6 \times 10^{-19} \approx -2000$ C, and there will be a charge of +2000 C on the positive ions. The force between the electrons and the rest of the wire is approximately given by the Coulomb force:

$$F = \frac{kq_1q_2}{r^2}$$

$$= \frac{9 \times 10^9 \times 2000^2}{(10^{-2})^2}$$

$$\approx 4 \times 10^{20} \text{ N}$$

This force is about the same as the gravitational force that holds the Moon to the Earth. Or put another way, it is the weight of over one hundred billion supertankers! The electrical force between the particles in a piece of wire is absolutely huge!

From this it can be seen that the electrical force on a moving electron in the wire in the magnetic field of another wire is a very delicate balance between two enormous forces—that from the positive protons in the wire and that from the negative electrons in the wire. Clearly any slight imbalance in those two forces will have an enormous effect. (You might like to calculate the force on just one electron from the positive charges in the wire in the previous example.) Although the Lorentz contraction is very slight, it is enough to produce a very small imbalance in the force on our electron. You can confirm that the Lorentz factor differs from 1 only in about the 23rd decimal place, but if that figure is multiplied by something like 10^{25} N, you end up with a normal sort of force; in fact, you end up with what is call the magnetic force.

So the magnetic force is actually a normal Coulomb force that results from a slight imbalance in the huge forces between all the protons and all the electrons in a wire in which many of the electrons are moving. Again you can see that relativity is actually a simplifying principle. What were thought to be two different, but related, forces are actually different aspects of the one electromagnetic force. The fact that all the electric motors used every day work so easily and efficiently, whether you are at home listening to a CD or flying around the world in an aircraft, is excellent evidence of the validity and relevance of Einstein's great theory.

7.5 Review

SUMMARY

- Relativistic momentum includes the Lorentz factor, γ, and hence, as more impulse is added, the mass seems to increase towards infinity as the speed gets closer, but never equal, to c. The relativistic momentum equation is:
  $p = \gamma mv = \gamma p_0$
- A term called relativistic mass, γm, may be used to indicate the mass of an object which is moving.
- Einstein found that the total energy of an object was given by:
  $E_{total} = E_k + E_{rest} = \gamma mc^2$
- The kinetic energy is given by:
  $E_k = (\gamma - 1)mc^2$
- The rest energy, which is the energy associated with the rest mass of an object, is given by:
  $E_{rest} = mc^2$
- Mass and energy are seen as different forms of the same thing. This means that mass, m, can be converted into energy, and energy can be converted into mass.
- Nuclear fission and fusion reactions result in a mass defect (change). It is this difference in mass that is converted to the energy released in nuclear reactions. This mass is related to the energy produced according to: $\Delta E = \Delta mc^2$.
- Nuclear fusion is the combining of light nuclei to form heavier nuclei. Extremely high temperatures are required for fusion to occur. This is the process occurring in stars.
- Hydrogen nuclei fuse to form deuterium. Further fusions result in the formation of isotopes of helium.

KEY QUESTIONS

1 Calculate the relativistic momentum of the *Rosetta* spacecraft as observed by the scientists at the European Space Agency. *Rosetta's* rest mass is 1230 kg and its speed was 775 m s^{-1}.

2 Calculate the relativistic momentum of a carbon-12 nucleus in a linear accelerator if its rest mass is $1.99264824 \times 10^{-26}$ kg and it is travelling at 0.850c.

3 Calculate the relativistic momentum of another carbon-12 nucleus in the solar wind if its rest mass is $1.99264824 \times 10^{-26}$ kg and it is travelling at a speed of 800 m s^{-1}.

The following information relates to questions 4–6.

A very fast arrow has a rest mass of 12.3 g and a speed of 0.750c.

4 Calculate the relativistic kinetic energy of the arrow.

5 Calculate the kinetic energy of the arrow according to the classical equation.

6 What accounts for the difference between the kinetic energy of the arrow in the relativistic calculation and the kinetic energy in the classical calculation?

A the difference in the arrow's velocity in the two calculations
B the difference in the arrow's momentum in the two calculations
C the difference in the arrow's rest mass in the two calculations
D the presence of the Lorentz factor in the relativistic calculation

7 Calculate the total energy of a very fast, vintage Vespa scooter if its rest mass is 210 kg and it is travelling at a speed of 2.55×10^8 m s^{-1}.

8 Calculate the energy produced by the Sun in one day if 4.00 million tonnes of matter are converted into energy every second.

9 The equation for the fusion of two isotopes of hydrogen (deuterium and tritium) is shown below.

$^2_1H + ^3_1H \rightarrow ^4_2He + ^1_0H$

Which one of the following best explains why energy is released during this process?

A Nucleons are created.
B Nucleons are lost.
C The nucleons lose mass.
D The nucleons gain mass.

10 What is the result of a large increase in the impulse provided to an object moving at a speed near that of light?

A a large change in the velocity of the object
B a proportional increase in the velocity of the object
C a very small increase in the velocity of the object
D no change in the velocity of the object at all

Chapter review

07

KEY TERMS

breaking point
collision
deformation
elastic
elastic collision
elastic limit
gravitational potential energy
impulse
inelastic collision
mechanical energy
momentum
nuclear fusion
strain potential energy
work

1 Which of the following statements correctly describe impulse? More than one correct answer is possible.

A Impulse is the rate of change of momentum.

B Impulse is the final momentum minus the initial momentum.

C Impulse is a scalar.

D Impulse can be calculated from the force and the time for which it acts.

2 A batsman blocks a 160 g cricket ball travelling towards him at 100 km h^{-1}. The ball leaves his bat at 20 km h^{-1}. Calculate the magnitude of the change in momentum of the ball.

3 A squash ball that is repeatedly hit against a wall during a game becomes hot. This is because:

A The racquet gives the ball kinetic energy.

B The impulse is positive.

C The collisions are perfectly elastic.

D Kinetic energy is not conserved in the collision.

4 Two identical bowling balls, each of mass 4.0 kg, move towards each other across a frictionless horizontal surface with equal speeds of 3.0 m s^{-1}. During the collision, 20 J of the initial kinetic energy is transformed into heat and sound. After the collision, the balls move in opposite directions away from each other.

a Is momentum conserved in this collision?

b Is this an elastic or inelastic collision? Explain your answer.

c Calculate the speed of each ball after the collision.

5 Calculate the magnitude of the average force required to be applied by the brakes of a 15 kg bicycle with a 65 kg rider if the bike and rider are travelling at 12 m s^{-1} and come to rest in 2.0 s.

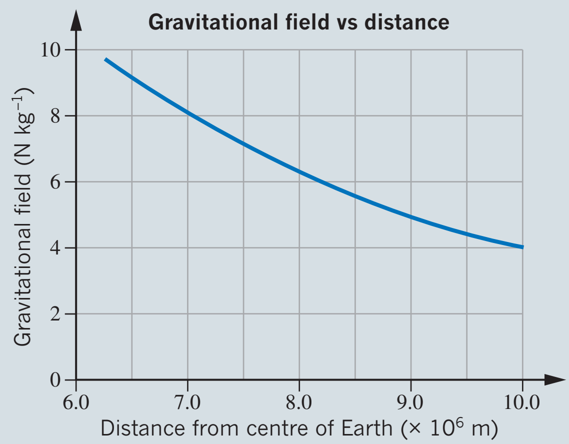
6 An arctic research worker uses a tractor to drag a sled with supplies across a glacier. The harness is held at an angle of 60° to the horizontal and applies a force of 300 N on the sled, which is initially at rest. A constant frictional force of 105 N acts on the sled as it is dragged for a distance of 240 m.

a For this distance, calculate the work done by the tractor on the 150 kg sled.

b Find the speed of the sled at the end of the 240 m stretch.

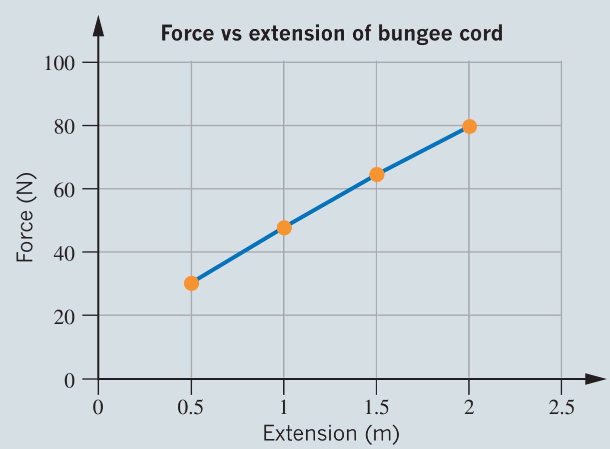
7 An 11 tonne satellite is in orbit at an altitude of 1100 km. A booster rocket is fired putting the satellite into an orbit at an altitude of 2100 km.

a Calculate the work done by the booster rocket to increase the potential energy of the satellite using the graph below. Assume the radius of Earth is 6.4×10^6 m.

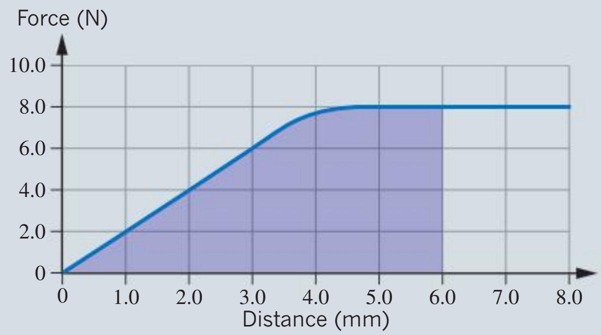


b Find the kinetic energy of the satellite in its final orbit.

8 Matthew decides that to increase his upper body strength he is going to stretch a piece of bungee cord 150 times each morning before school, grasping one end in each hand. If the force–distance curve is given below, and he stretches the cord out from 0.5 m to 1 m, estimate how much energy is expended in his workout.



9 A steel cable of length 1.50 m is stretched by fixing it at one end, and applying a force to the other end. The graph of the force applied and the extension is shown below:



Calculate the strain energy stored in the cable when stretched by a distance of 6.0 mm.

10 Two children are standing on a bridge throwing stones into a river below. Susan throws a stone upwards, and Peter throws his downwards at the same speed. Select the correct answer.

A Both stones will hit the water at the same speed.

B The stone that is thrown downwards by Peter will hit the water at a greater speed than Susan's stone which was thrown upwards.

C Susan's stone will hit the water at a greater speed than Peter's stone.

D More information is required to determine which stone hits the water at the greatest speed.

The following information relates to questions 11–14.

A 200 g toy truck with a springy bumper travelling at 0.300 m s^{-1} collides with a 100 g toy car travelling in the same direction at 0.200 m s^{-1}.

The car moves forward travelling at an increased speed of 0.300 s^{-1}.

11 Calculate the speed of the truck after the collision.

12 Calculate the total kinetic energy of the system before the collision.

13 Calculate the total kinetic energy of the system after the collision.

14 Complete the following statements by selecting the appropriate option from those in bold.

a The total kinetic energy before the collision is **more than**/**less than**/**equal to** the total kinetic energy after the collision.

b The kinetic energy of the system of toys **is**/**is not** conserved.

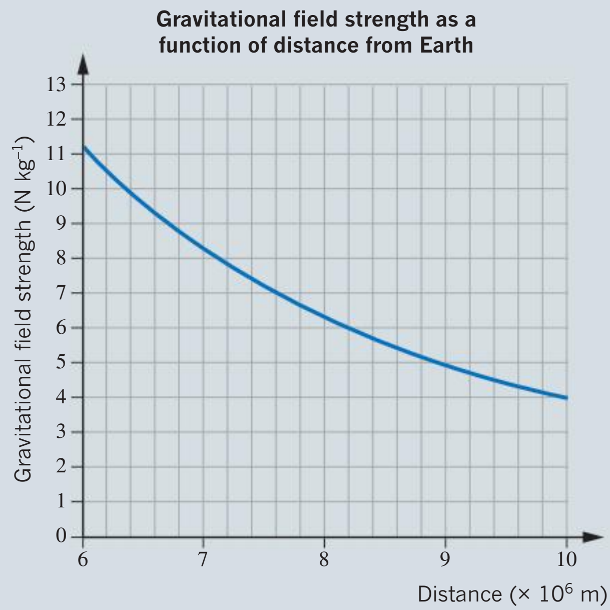
c The total energy of the system of toys **is**/**is not** conserved.

d The total momentum of the system of toys **is**/**is not** conserved.

e The collision **is/is not** perfectly elastic because **kinetic energy**/**total energy**/**momentum** is not conserved.

15 An 80 kg student jumps from a bridge on a bungee rope. If the 100 m rope stretches by 10%, calculate the spring constant of the rope.

16 A new space telescope is 500 km above the surface of Earth in a circular orbit. Use the graph below to calculate its gravitational potential energy relative to the surface of Earth if the mass of the telescope is 11×10^6 kg.



17 If a spaceship is travelling at 99% of the speed of light, which of the following best explains why it can't simply turn on its engine and accelerate through and beyond the speed of light, *c*, as the increase in momentum should be equal to the impulse applied?

A The law of impulse equals change in momentum does not apply at speeds close to *c*.

B While the momentum increases with the impulse, it is the mass rather than the speed that is getting greater.

C The spaceship does actually exceed *c*, but it doesn't appear to from another frame of reference because of length contraction of the distance it covers.

D Given enough impulse the spaceship could exceed *c*, but no real spaceship could carry enough fuel.

18 Find the speed of a proton if it has kinetic energy equal to its rest mass energy.

19 Find the relativistic mass of the proton described in the question above, if $m_p = 1.67 \times 10^{-27}$ kg.

20 Calculate the relativistic kinetic energy of a bus with a rest mass of 5.30 tonnes and travelling at a speed of 0.960*c*.


T··Home

REVIEW QUESTIONS

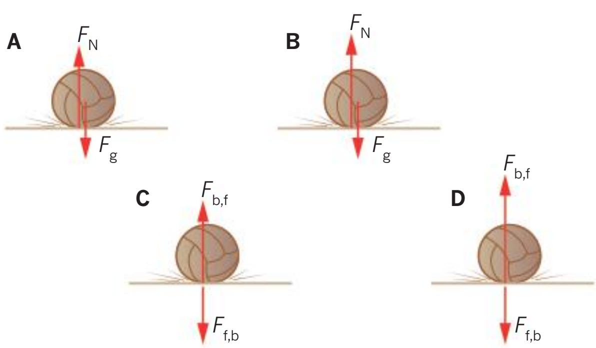
How fast can things go?

The following information relates to questions 1–2.

A netball is dropped vertically from a height of 1.5 m onto a horizontal floor. The diagrams below relate to the instant that the ball reaches the floor and is stationary for a short period of time before rebounding.

1 On a copy of the diagram below, draw and identify the forces that are acting on the ball at this instant, being careful to show the relative sizes of the forces.

2 Which of the following correctly represents the action/reaction forces acting between the ball and the floor at this instant? (More than one answer may be correct.).



The following information relates to questions 3–7.

In the Gravitron ride, the patrons enter a cylindrical chamber which rotates rapidly, causing them to be pinned to the vertical walls as the floor drops away. A particular Gravitron ride has a radius of 5.00 m and rotates with a period of 2.50 s. Jodie, of mass 60.0 kg, is on the ride.

3 Choose the correct responses in the following statement from the options given in bold:
As the Gravitron spins at a uniform rate and Jodie is pinned to the wall, the horizontal forces acting on her are **balanced**/**unbalanced** and the vertical forces are **balanced**/**unbalanced**.

4 Calculate the speed of Jodie as she revolves on the ride.

5 What is the magnitude of her centripetal acceleration?

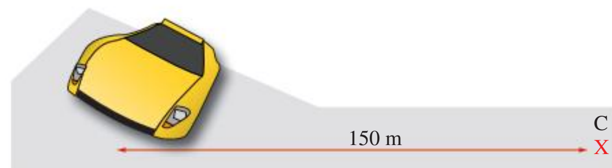
6 Calculate the magnitude of the normal force that acts on Jodie from the wall of the Gravitron.

7 The rate of rotation of the ride is increased so that Jodie completes six revolutions every 10.0 s. What is the frequency of Jodie's motion now?

The following information relates to questions 8–9.

A car racing track is banked so that when the cars corner at 40 m s^{-1}, they experience no sideways frictional forces. The track is circular with a radius of 150 m.

8 In the diagram below, the car is travelling at 40 m s^{-1}. Draw and identify the forces that are acting on the car in the vertical plane at this instant.



9 Calculate the angle to the horizontal at which the track is banked.

The following information relates to questions 10–12.

A skateboarder of mass 55 kg is practising on a half-pipe of radius 2.0 m. At the lowest point of the half-pipe, the speed of the skater is 6.0 m s^{-1}.

10 What is the acceleration of the skater at this point?

11 Calculate the size of the normal force acting on the skater at this point.

12 Describe the apparent weight of the skater as they travel through the lowest point in the pipe.

The following information relates to questions 13–16.

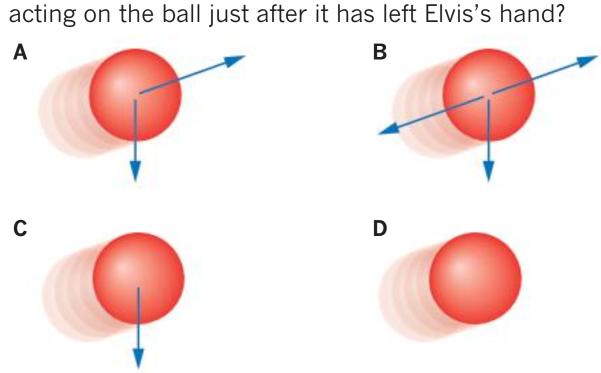
Two friends, Elvis and Kurt, are having a game of catch. Elvis throws a baseball to Kurt, who is standing 8.0 m away. Kurt catches the ball at the same height, 2.0 s after it is thrown. The mass of the baseball is 250 g. Ignore the effects of air resistance.

13 Determine the value of the maximum height gained by the ball during its flight.

14 What was the acceleration of the ball at its maximum height?

15 Calculate the speed at which the ball was thrown.

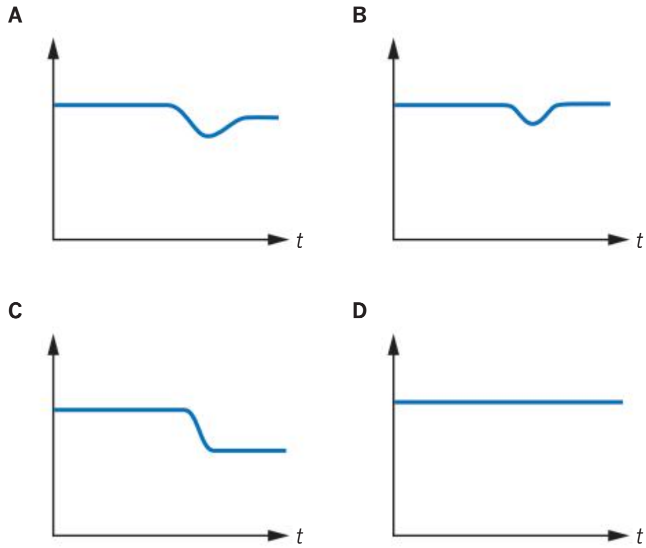
16 Which of the following diagrams best shows the forces acting on the ball just after it has left Elvis's hand?



UNIT 3 • Area of Study 3

The following information relates to questions 17–20.

A tennis ball is thrown at a stationary bowling ball of mass 5.0 kg. The tennis ball rebounds and the bowling ball rolls forward very slowly. The collision is considered to be inelastic. Use the graphs to answer questions 17 and 18.



17 Which graph best shows the total kinetic energy of the system before, during and after the collision? Explain your answer.

18 Which graph best shows the total momentum of the system before, during and after the collision? Explain your answer.

19 How does the change in momentum of the tennis ball compare with the change in momentum of the bowling ball?

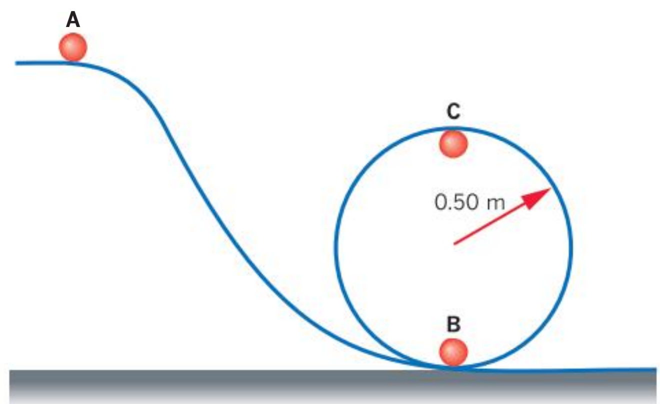
- **A** They are equal.
- **B** The tennis ball experiences a greater change of momentum.
- **C** The bowling ball experiences a greater change of momentum.
- **D** They are equal in magnitude and opposite in direction.

20 How do the forces that the two balls exert on each other during the collision compare?

- **A** The forces are equal.
- **B** The tennis ball exerts the greater force.
- **C** The bowling ball exerts the greater force.
- **D** The forces are equal in magnitude and opposite in direction.

The following information relates to questions 21–24.

A ball bearing of mass 25 g is rolled along a smooth track in the shape of a loop-the-loop. The ball bearing is given a launch speed at A so that it just maintains contact with the track as it passes through point C. Ignore drag forces when answering these questions.



21 Determine the magnitude of the acceleration of the ball bearing as it passes point C.

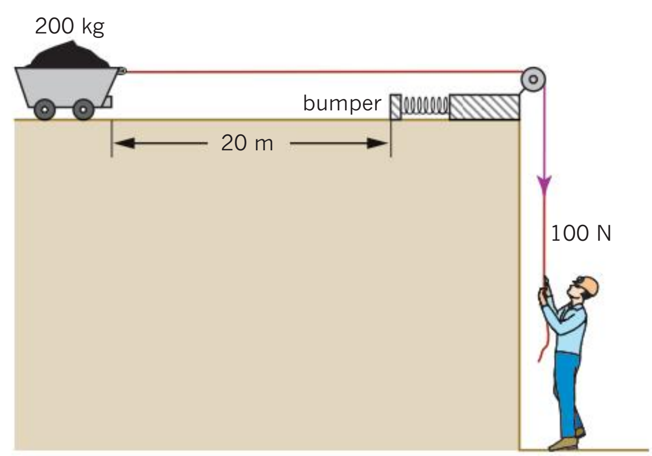
22 How fast is the ball bearing travelling at point C?

23 What is the apparent weight of the ball bearing at point C?

24 How fast is the ball bearing travelling at point B?

The following information relates to questions 25–29.

A small-time gold prospector sets up a cable-pulley system that allows him to move a container full of ore of total mass 200 kg from rest a distance of 20 m along a level section of rail track, as shown in the following diagram.



When the load reaches the end of the track, it is momentarily brought to rest by a powerful spring-bumper system, which is assumed to have negligible mass.

A constant frictional force of 30 N acts on the wheels of the container along the track. Assume that there is negligible friction between the pulley and the cable.

The prospector applies a constant force of 100 N to the rope as the trolley moves along the track.

25 How much work is done on the container as it moves along the track?

26 Calculate the change in kinetic energy of the load as it moves along the track.

27 What is the speed of the load when it reaches the end of the track?

28 What is the power output of the prospector as he moves one 200 kg load over the track during a 10 s interval?

29 The spring bumper has a force constant of 1500 N m^{-1}. How much kinetic energy does the container lose as the spring compresses by 18 cm?

The following information relates to questions 30–33.

At football training, some of the players are throwing themselves at a large tackle bag of mass 45 kg. During one exercise, a ruckman of mass 120 kg running at 6.0 m s^{-1} crashes into the stationary bag and carries it forward.

30 What is the combined speed of the bag and ruckman?

31 How much momentum does the ruckman lose?

32 How much momentum does the tackle bag gain?

33 Is the collision of the ruckman and the tackle bag elastic or inelastic? Use calculations to justify your answer.

The following information relates to questions 34–38.

A physics student decides to study the properties of a bungee rope by recording the extension produced by various masses attached to the end of a section of the rope. The results of the experiment are shown in the following table.

Mass (kg)	Extension (m)
0.5	0.24
1.0	0.52
1.5	0.73
2.0	0.95
2.5	1.20
3.0	1.48
3.5	1.70

34 Draw the force versus extension graph for this bungee rope.

35 Estimate the value of the spring constant for this rope.

During an investigation, the student stretched the rope horizontally by 15 m.

36 Assuming that the rope behaves ideally, determine the potential energy stored in the bungee rope at this point.

Finally, the student stands on a skateboard and allows the rope, stretched by 15 m, to drag her across the smooth floor of the school gymnasium.

37 Which statement best describes the motion of the student?

A She moves with a constant velocity.
B She moves with a constant acceleration.
C She moves with increasing velocity and decreasing acceleration.
D She moves with increasing velocity and increasing acceleration.

38 The combined mass of the student and her board is 60 kg. Calculate the maximum speed that she attains as she is pulled by the bungee cord.

39 In 1905 Einstein put forward two postulates. Which two of the following best summarise them?

A All observers will find the speed of light to be the same.
B In the absence of a force, motion continues with constant velocity.
C There is no way to detect an absolute zero of velocity.
D Absolute velocity can only be measured relative to the aether.

40 You are in interstellar space and know that your velocity relative to Earth is 4×10^6 m s^{-1} away from it. You then notice another spacecraft with a velocity, towards you, of 4×10^5 m s^{-1}. Which one or more of the following best describes the velocity of the other craft?

A Away from Earth at 3.6×10^6 m s^{-1}
B Towards Earth at 3.6×10^6 m s^{-1}
C Away from Earth at 4.4×10^6 m s^{-1}
D Towards Earth at 4.4×10^6 m s^{-1}

41 One of the fastest objects made on Earth was the Galileo probe which, as a result of Jupiter's huge gravity, entered its atmosphere in 1995 at a speed of nearly 50 000 m s^{-1}. Which of the following is the best estimate of the Lorentz factor for the probe?

A Less than 1
B 1.000 000 00
C 1.000 000 01
D 1.1

42 Which one of the following best represents the basis of Einstein's considerations, which eventually led to the theory of special relativity?

A The results of numerous experiments to determine the speed of light.
B The work of Isaac Newton and Michael Faraday.
C His consideration of the consequences of accepting the implications of Maxwell's equations.
D His own experiments in electromagnetism.

43 Aristotle suggested that the 'natural' state of motion for any object is rest. Galileo introduced the principle of inertia, which suggested that the natural state of motion is constant velocity (zero velocity being just one example). Explain why Aristotle's view was so hard to shake, and why, if we had spent time as an astronaut in a space station, Galileo's principle would be much easier to accept.

44 You are in a spaceship travelling at very high speed past a new colony on Mars. Do you notice time going slowly for you; for example, do you find your heart rate is slower than normal? Do the people on Mars appear to be moving normally? Explain your answers.

The following information relates to questions 45–47.

The star Xquar is at a distance of 5 light-years from Earth. Space adventurer Raqu heads from Earth towards Xquar at a speed of 0.9c.

45 For those watching from Earth, how long will it take for Raqu to reach Xquar?

46 From Raqu's point of view how long will it take her to reach Xquar?

47 Explain why it is that, although Raqu knew that Xquar was 5 light-years from Earth, and that she was to travel at 0.9c, it took much less time than might be expected from these figures.

The following information relates to questions 48–50.

The fusion reaction that powers the Sun effectively combines four protons (rest mass 1.673×10^{-27} kg) to form a helium nucleus of two protons and two neutrons (total rest mass 6.645×10^{-27} kg). The total power output of the Sun is a huge 3.9×10^{26} W.

48 How much energy is released by each fusion of a helium nucleus?

49 How many helium nuclei are being formed every second in the Sun?

50 How much mass is the Sun losing every day?

UNIT 4 How can two contradictory models explain both light and matter?

AREA OF STUDY 1

How can waves explain the behaviour of light?

Outcome 1: On completion of this unit the student should be able to apply wave concepts to analyse, interpret and explain the behaviour of light.

AREA OF STUDY 2

How are light and matter similar?

Outcome 2: On completion of this unit the student should be able to provide evidence for the nature of light and matter, and analyse the data from experiments that supports this evidence.

AREA OF STUDY 3

Practical investigation

Outcome 3: On completion of this unit the student should be able to design and undertake a practical investigation related to waves or fields or motion, and present methodologies, findings and conclusions in a scientific poster.

To achieve this outcome the student will draw on key knowledge outlined in Area of Study 3 and the related key science skills.

CHAPTER

08 Properties of mechanical waves

Have you ever watched ocean waves heading toward the shore? For many people their first thought when encountering a topic called 'waves' is to picture a water wave moving across the surface of an ocean. The wave may be created by some kind of disturbance like the action of wind on water or a boat as it moves through the water.

Waves are, in fact, everywhere. Sound, visible light, radio waves, waves in the string of an instrument, the wave of a hand, the 'Mexican wave' at a stadium and the recently discovered gravitational waves—all are waves or wave-like phenomena. Understanding the physics of waves provides a broad base upon which to build your understanding of the physical world. A knowledge of waves gives an introduction to the concepts that describe the nature of light.

Key knowledge

By the end of this chapter you will have studied the properties of mechanical waves, and will be able to:

- explain a wave as the transmission of energy through a medium without the net transfer of matter
- distinguish between transverse and longitudinal waves
- identify the amplitude, wavelength, period and frequency of waves
- calculate the wavelength, frequency, period and speed of travel of waves using: $v = f\lambda$, and $\frac{\lambda}{T}$
- explain qualitatively the Doppler effect
- explain resonance as the superposition of a travelling wave and its reflection, and with reference to a forced oscillation matching the natural frequency of vibration
- analyse the formation of standing waves in strings fixed at one or both ends.

8.1 Longitudinal and transverse waves

Throw a stone into a pool or lake, and you will see circular waves form and move outwards from the source as ripples, as shown in Figure 8.1.1. Stretch a cord out on a table and wriggle one end back and forth across the table surface and another type of wave can be observed. Sound waves, water waves and waves in strings are all examples of **mechanical waves**. Mechanical waves, as opposed to electromagnetic waves, cannot transmit energy through a vacuum. Mechanical waves are the focus of this chapter.

FIGURE 8.1.1 The ripples in a pond indicate a transfer of energy.

MECHANICAL WAVES

Watch a piece of driftwood, a leaf, or even a surfer, resting in the water as a smooth wave goes past. The object moves up and down but doesn't move forward with the wave. The movement of the object on the water reveals how the particles in the water move as the wave passes; that is, the particles in the water move up and down from an average position.

Any wave that needs a **medium** (such as water) through which to travel is called a mechanical wave. Mechanical waves can move over very large distances but the particles of the medium only have very limited movement.

Mechanical waves transfer energy from one place to another through a medium. The particles of the matter vibrate back and forth or up and down about an average position, which transfers the energy from one place to another. For example, energy is given to an ocean wave by the action of the wind far out at sea. The *energy* is transported by waves to the shore but (except in the case of a tsunami event) most of the ocean water itself does not travel onto the shore.

A wave involves the transfer of energy without the net transfer of matter.

PULSES VERSUS PERIODIC WAVES

A single wave **pulse** can be formed by giving a slinky spring or rope a single up and down motion as shown in Figure 8.1.2(a). As the hand pulls upwards, the adjacent parts of the slinky will also feel an upward force and begin to move upward. The source of the wave energy is the movement of the hand.

If the up and down motion is repeated, each successive section of the slinky will move up and down, moving the wave forward along the slinky as shown in Figure 8.1.2(b). Connections between each loop of the slinky cause the wave to travel away from the source, carrying with it the energy from the source.

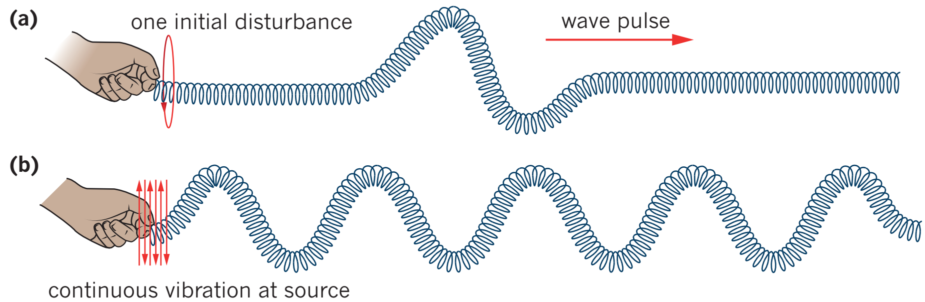


FIGURE 8.1.2 (a) A single wave pulse can be sent along a slinky by a single up and down motion. (b) A continuous or periodic wave is created by a regular, repeated movement of the hand.

In a continuous wave or periodic wave, continuous vibration of the source, such as that shown in Figure 8.1.2(b), will cause the particles within the medium to oscillate about their average position in a regular, repetitive or periodic pattern. The source of any mechanical wave is this repeated motion or vibration. The energy from the vibration moves through the medium and constitutes a mechanical wave.

Transverse waves

When waves travel on water, or through a rope, spring or string, the particles within the medium vibrate up and down in a direction perpendicular, or **transverse**, to the direction of motion of the wave energy (see Figure 8.1.3). Such a wave is called a *transverse wave*. When the particles are displaced upwards from the average position, or resting position, they reach a maximum positive displacement called a **crest**. Particles below the average position fall to a maximum negative position called a **trough**.

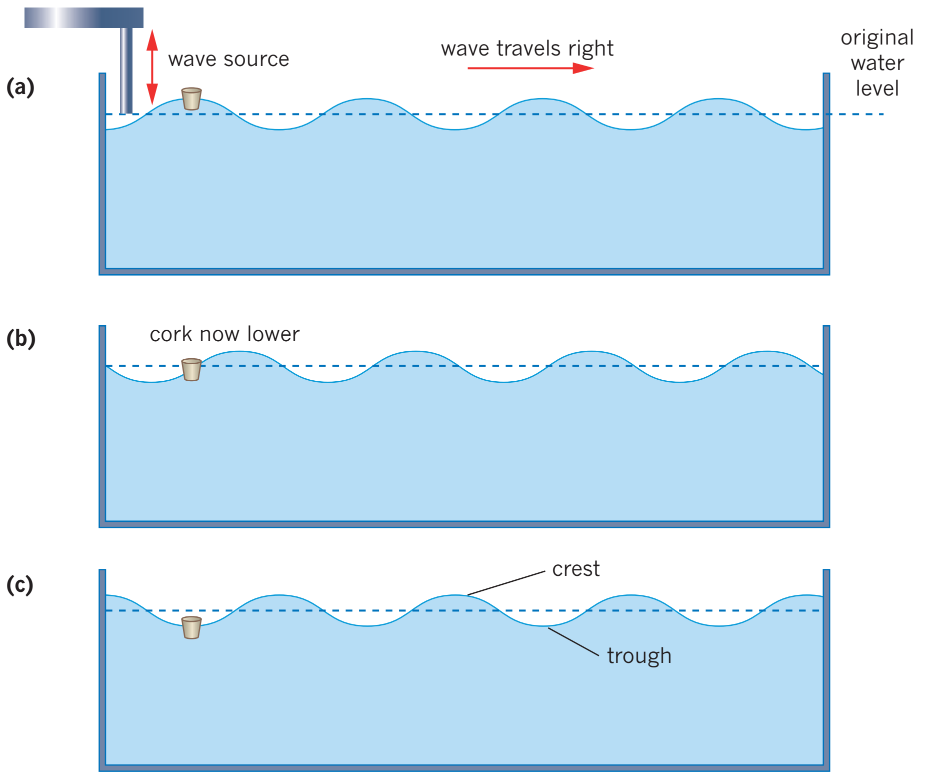


FIGURE 8.1.3 A continuous water wave moves to the right. As it does so the up and down displacement of the particles transverse to the wave motion can be monitored using a cork. The cork simply moves up and down as the wave passes through it.

Longitudinal waves

In a **longitudinal** mechanical wave, the vibration of the particles within the medium are in the same direction, or parallel to, the direction of energy flow of the wave. You can demonstrate this type of wave with a slinky by moving your hand backwards and forwards in a line parallel to the length of the slinky, as shown in Figure 8.1.4.

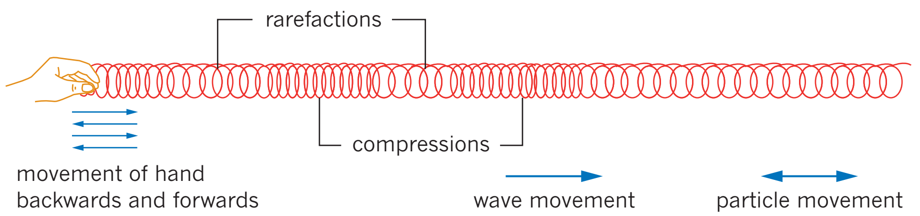


FIGURE 8.1.4 When the direction of the vibrations of the medium and the direction of travel of the wave energy are parallel, a longitudinal wave is created. This can be demonstrated with a slinky.

As you move your hand, a series of compressed and expanded areas form along the slinky. **Compressions** are those areas where the coils of the slinky come together. Expansions are regions where the coils are spread apart. Areas of expansion are termed **rarefactions**. The compressions and rarefactions in a longitudinal wave correspond to the crests and troughs of a transverse wave.

An important example of a longitudinal wave is a sound wave. As the cone of a loudspeaker vibrates, the layer of air next to it is alternately pushed away and drawn back creating a series of compressions and rarefactions in the air (see Figure 8.1.5). This vibration is transmitted through the air as a sound wave. Like transverse waves, the individual molecules vibrate over a very small distance while the wave itself can carry energy over very long distances. If the vibration was from a single point then the waves would tend to spread out spherically.

FIGURE 8.1.5 The motion of a flame in front of a loudspeaker is clear evidence of the continuous movement of air backwards and forwards as the loudspeaker creates a sound wave.

PHYSICSFILE

Water waves

Water waves are often classified as transverse waves, but this is an approximation. In practical situations, transverse and longitudinal waves don't always occur in isolation. The breaking of waves on a beach produces complex wave forms which are a combination of transverse and longitudinal waves (see Figure 8.1.6).

If you looked carefully at a cork bobbing about in gentle water waves you would notice that it doesn't move straight up and down but that it has a more elliptical motion. It moves up and down, and very slightly forwards and backwards as each wave passes. However, since this second aspect of the motion is so subtle, in most circumstances it is adequate to treat water waves as if they were purely transverse waves.

FIGURE 8.1.6 Even though this surfer rides forward on the wave, the water itself only moves in an elliptical motion as the wave passes.

8.1 Review

SUMMARY

- Vibrating objects transfer energy through waves, travelling outwards from the source. Waves on water, on a string and sound waves in air are examples of waves.
- A wave may be a single pulse or it may be continuous or periodic (successive crests and troughs or compressions and rarefactions).
- A wave only transfers energy from one point to another. There is no net transfer of matter or material.
- Mechanical waves can be either transverse or longitudinal.
- In a transverse wave, the oscillations are perpendicular to the direction in which the wave energy is travelling. A wave in a string is an example of a transverse wave.
- In a longitudinal wave, the oscillations are parallel to (along) the direction the wave energy is travelling. Sound is an example of a longitudinal wave.

KEY QUESTIONS

1 Describe the motion of particles within a medium as a mechanical wave passes through the medium.

2 Which of the following statements are true and which are false? For the false statements, rewrite them so they become true.

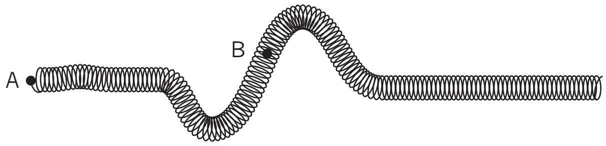
a Longitudinal waves occur when particles of the medium vibrate in the opposite direction to the direction of the wave.

b Transverse waves are created when the direction of vibration of the particles is at right angles to the direction of the wave.

c A longitudinal wave is able to travel through air.

d The vibrating string of a guitar is an example of a transverse wave.

3 The diagram below represents a slinky spring held at point A by a student.



Draw an image of the pulse a short time after that shown in the diagram and determine the motion of point B. Is point B moving upwards, downwards or is it stationary?

4 Which of the following are examples of mechanical waves?

light, sound, ripples on a pond, vibrations in a rope

5 The diagram below shows dots representing the average displacement of air particles at one moment in time as a sound wave travels to the right.

Describe how particles A and B have moved from their equally-spaced undisturbed positions to form the compression.

6 A sound wave is emitted from a speaker and heard by Lee who is 50 m from the speaker. He made several statements once he heard the sound. Which one or more of the following statements made by Lee would be correct? Explain your answers.

A Hearing a sound wave tells me that air particles have travelled from the speaker to me.

B Air particles carried energy with them as they travelled from the speaker to me.

C Energy has been transferred from the speaker to me.

D Energy has been transferred from the speaker to me by the oscillation of air particles.

7 A mechanical wave may be described as transverse or longitudinal. In a *transverse* wave, how does the motion of the particles compare with the direction of travel of the wave?

8 Classify the waves described below as either longitudinal or transverse:

a sound waves

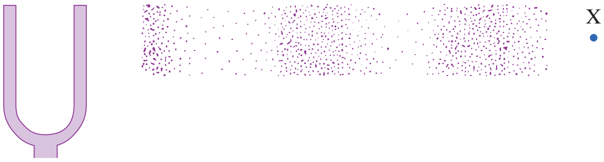
b a vibrating guitar string

c slinky moved with an upward pulse

d slinky pushed forwards and backwards.

9 Mechanical waves generally travel faster in solids than in gases. Provide an explanation for this.

10 For the wave shown below, describe the direction of energy transfer of the sound between the tuning fork and point X. Justify your answer.



8.2 Measuring mechanical waves

The features of a mechanical wave can be represented using a graph. In this section you will explore how the displacement of particles within the wave can be represented using graphs. From these graphs several key features of a wave can be identified:

- amplitude
- wavelength
- frequency
- period
- speed.

Waves of different amplitudes and wavelengths can be seen in Figure 8.2.1.

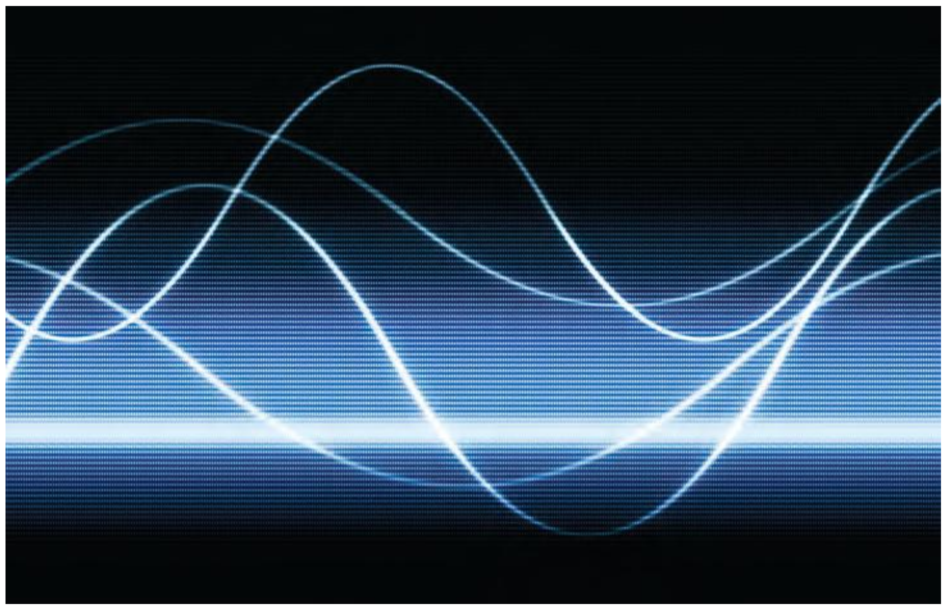


FIGURE 8.2.1 Waves can have different wavelengths, amplitudes, frequencies, periods and velocities, which can all be represented on a graph.

DISPLACEMENT–DISTANCE GRAPHS

The displacement–distance graph in Figure 8.2.2 shows the displacement of all particles along the length of a transverse wave at a particular point in time.

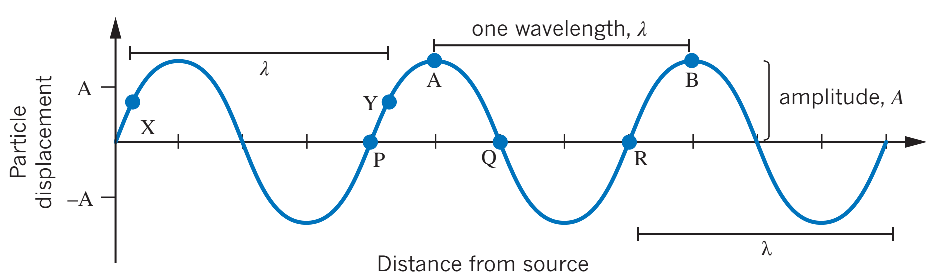


FIGURE 8.2.2 A sine wave representing the particle displacements along a wave.

Have a look back at Figure 8.1.2(b) on page 275 of a continuous wave in a slinky. This 'snapshot' in time shows the particles moving up and down **sinusoidally** about a central rest position. As a wave passes a given point, the particle at that point will go through a complete cycle before returning to its starting point. The wave spread along the length of the slinky has the shape of a sine or cosine function, which you will recognise from mathematics. A displacement–distance graph shows the position (displacement) of the particles at any moment in time along the slinky about a central position.

From a displacement–distance graph, the amplitude and wavelength of a wave are easily recognisable.

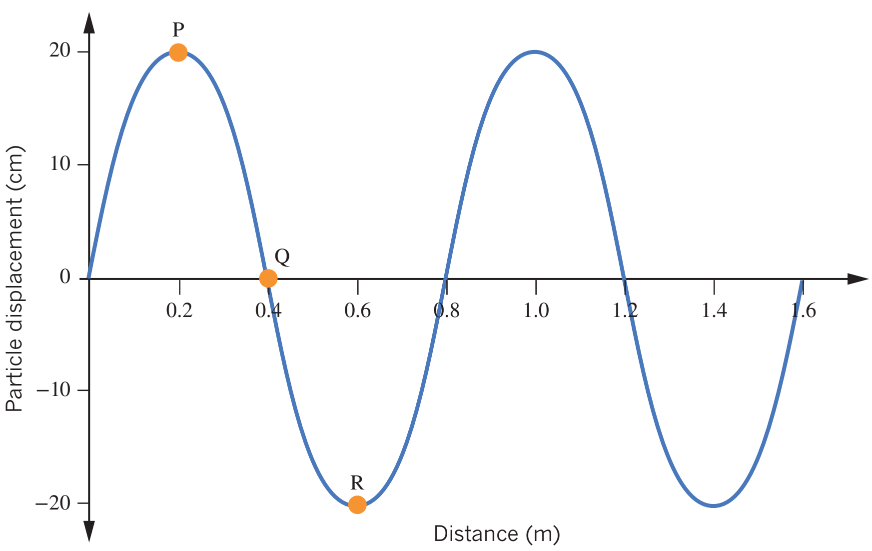
- The **amplitude** of a wave is the maximum displacement of a particle from the average or rest position. That is, the amplitude is distance from the middle of a wave to the top of a crest or to the bottom of a trough. The total distance a particle will move through in one cycle is twice the amplitude.
- The **wavelength** of a wave is the distance between any two successive points in phase (e.g. points A and B or X and Y in Figure 8.2.2). It is denoted by the Greek letter λ (lambda), and is measured in metres. Two particles on the wave are said to be in **phase** if they have the same displacements from the average position and are moving in the same direction. Points P and R in Figure 8.2.2 are two such particles that are in phase, as are points A and B and X and Y but not P and Q.

The **frequency**, *f*, is the number of complete cycles that pass a given point per second and is measured in hertz (Hz). By drawing a series of displacement–distance graphs at various times, you can see the motion of the wave. By comparing the changes in these graphs, the travelling speed and direction of the wave can be found, as well as the direction of motion of the vibrating particles.

Worked example 8.2.1

DISPLACEMENT–DISTANCE GRAPH

The displacement–distance graph below shows a snapshot of a transverse wave as it travels along a spring towards the right. Use the graph to determine the amplitude and the wavelength of this wave.

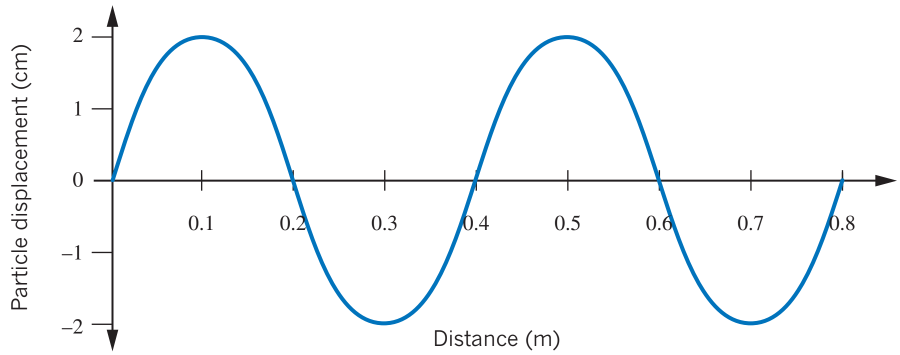


Thinking	Working
Amplitude on a displacement–distance graph is the distance from the average position to a crest (P) or a trough (R). Read the displacement of a crest or a trough from the vertical axis. Convert to SI units where necessary.	Amplitude = 20 cm = 0.2 m
Wavelength is the distance for one complete cycle. Any two consecutive points in phase and at the same position on the wave could be used.	The first cycle runs from the origin through P, Q, R to intersect the horizontal axis at 0.8 m. This intersection is the wavelength. Wavelength λ = 0.8 m

Worked example: Try yourself 8.2.1

DISPLACEMENT–DISTANCE GRAPH

The displacement–distance graph below shows a snapshot of a transverse wave as it travels along a spring towards the right. Use the graph to determine the wavelength and the amplitude of this wave.



DISPLACEMENT–TIME GRAPHS

A displacement–time graph such as the one shown in Figure 8.2.3 tracks the position of one point over time as the wave moves through that point.

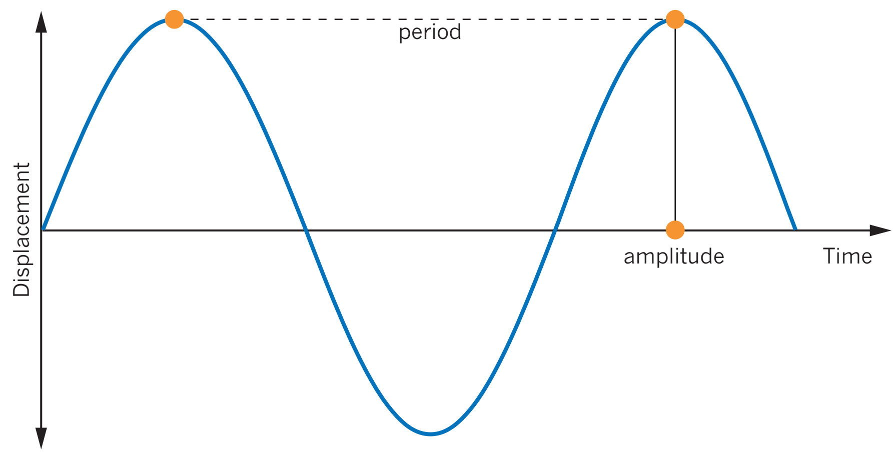


FIGURE 8.2.3 The graph of displacement versus time from the source of a transverse wave shows the movement of a *single point* on a wave over time as the wave passes through that point.

The displacement–time graph looks very similar to a displacement–distance graph of a transverse wave, so be careful to check the horizontal axis label.

Crests and troughs are shown the same way in both graphs. The amplitude is still the maximum displacement from the average or rest position of either a crest or a trough. But the distance between two successive points in phase in a displacement–time graph represents the **period** of the wave, T, measured in seconds.

The period is the time it takes for any point on the wave to go through one complete cycle (e.g. from crest to successive crest). The period of a wave is inversely related to its frequency:

$$T = \frac{1}{f}$$

where T is the period of the wave (s)

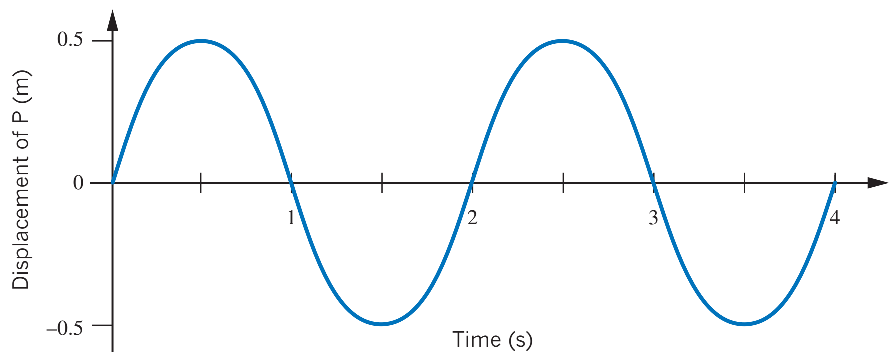
f is the frequency of the wave (Hz)

The amplitude and period of a wave, and the direction of motion of a particular particle, can be determined from a displacement–time graph.

Worked example 8.2.2

DISPLACEMENT–TIME GRAPHS

The displacement–time graph below shows the motion of a single part of a rope (point P) as a wave passes by travelling to the right. Use the graph to find the amplitude, period and frequency of the wave.

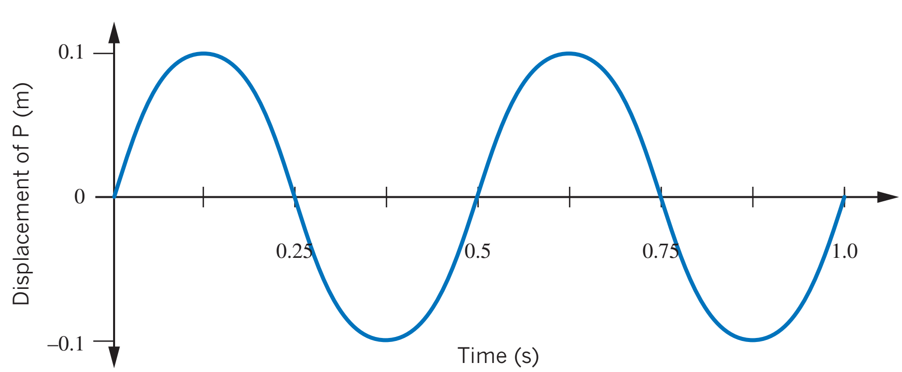


Thinking	Working
The amplitude on a displacement–time graph is the displacement from the average position to a crest or trough. Note the displacement of successive crests and/or troughs on the wave and carefully note units on the vertical axis.	Maximum displacement is 0.5 m Therefore amplitude = 0.5 m
Period is the time it takes to complete one cycle and can be identified on a displacement–time graph as the time between two successive points that are in phase. Identify two points on the graph at the same position in the wave cycle, e.g. the origin and t = 2 s. Confirm by checking two other points, e.g. two crests or two troughs.	Period T = 2 s
Frequency can be calculated using $f = \frac{1}{T}$, measured in hertz (Hz).	$f = \frac{1}{T} = \frac{1}{2}$ = 0.5 Hz

Worked example: Try yourself 8.2.2

DISPLACEMENT–TIME GRAPHS

The displacement–time graph below shows the motion of a single part of a rope as a wave passes travelling to the right. Use the graph to find the amplitude, period and frequency of the wave.



THE WAVE EQUATION

Although the speed of a wave can vary, there is a relationship between the speed of a wave and other significant wave characteristics.

Think back to the study of motion. Speed is given by:

$$v = \frac{\text{distance travelled}}{\text{time taken}} = \frac{d}{\Delta t}$$

This can be rewritten in terms of the distance of one wavelength (λ) in one period (T), which will be:

$$v = \frac{\lambda}{T}$$

and since

$$f = \frac{1}{T}$$

the relationship becomes:

$$v = f\lambda$$

where v is the speed (m s^{-1})

f is the frequency (Hz)

λ is the wavelength (m)

This is known as the wave equation and applies to both longitudinal and transverse mechanical waves.

Worked example 8.2.3

THE WAVE EQUATION

A longitudinal wave has a wavelength of 2.0 m and a speed of 340 m s^{-1}. What is the frequency, f, of the wave?	
Thinking	**Working**
The wave equation states that $v = f\lambda$. Knowing both v and λ, the frequency, f, can be found. Rewrite the wave equation in terms of f.	$v = f\lambda$ $f = \frac{v}{\lambda}$
Substitute the known values and solve.	$f = \frac{v}{\lambda}$ $= \frac{340}{2.0}$ $= 170$ Hz

Worked example: Try yourself 8.2.3

THE WAVE EQUATION

A longitudinal wave has a wavelength of 4.0×10^{-7} m and a speed of 3.0×10^{8} m s^{-1}. What is the frequency, f, of the wave?

Worked example 8.2.4

THE WAVE EQUATION

A longitudinal wave has a wavelength of 2.0 m and a speed of 340 m s^{-1}. What is the period, *T*, of the wave?	
Thinking	**Working**
Rewrite the wave equation in terms of *T*.	$v = f\lambda$, and $f = \frac{1}{T}$ $v = \frac{1}{T} \times \lambda$ $= \frac{\lambda}{T}$ $T = \frac{\lambda}{v}$
Substitute the known values and solve.	$T = \frac{\lambda}{v}$ $= \frac{2.0}{340}$ $= 5.9 \times 10^{-3}$ s

Worked example: Try yourself 8.2.4

THE WAVE EQUATION

A longitudinal wave has a wavelength of 4.0×10^{-7} m and a speed of 3.0×10^{8} m s^{-1}. What is the period, *T*, of the wave?

THE DOPPLER EFFECT

The **Doppler effect** is a phenomenon of waves that is observed whenever there is relative movement between the source of the waves and an observer. It causes an apparent *increase* in frequency when the relative movement is *towards* the observer (i.e. the distance between observer and wave source is *decreasing*) and an apparent *decrease* in frequency when the relative movement is *away* from the observer (i.e. the distance between observer and wave source is *increasing*). It can be observed for any type of wave and has been particularly useful in astronomy for understanding the expanding universe.

Named after Austrian physicist Christian Doppler, who proposed it in 1842, the Doppler effect only affects the *apparent* frequency of the wave. The actual frequency of the wave does not change. A common experience of the Doppler effect is in listening to the sound of a siren from an emergency vehicle as it approaches and passes by.

Suppose a wave source, such as an ambulance siren, is stationary relative to an observer. The observer will receive and hear the disturbances (rarefactions and compressions in this example) at the same rate as the source creates them. If the wave source were to travel towards the observer, then each consecutive disturbance will originate from a position a little closer than the previous one. Hence each disturbance will have a little less distance to travel before reaching the observer than the one immediately before it. The effect is that the frequency of arrival of the disturbances is higher than the originating frequency (see the person on the right in Figure 8.2.4).

Alternatively, if the source is moving away from the observer, each consecutive disturbance will originate from a distance a little further away than the one immediately before and so has a greater distance to travel. The disturbances will arrive at the observer with a frequency that is less than the originating frequency (see the person on the left in Figure 8.2.4).

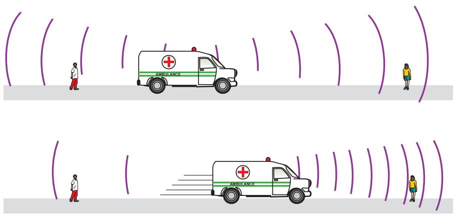


FIGURE 8.2.4 The Doppler effect. An object emitting a sound moving towards an observer (on the right) will emit sound waves closer together in its direction of travel and hence a higher frequency is heard by the observer. When the object is moving away from the observer (on the left), the sound waves are emitted further apart and hence a lower frequency is heard by the observer.

The net effect is that when the wave is moving towards an observer, the frequency of arrival of the wave will be higher than the frequency of the original source. When the wave is moving away from the observer, the frequency of arrival will be lower than the frequency of the original source. Therefore, as a result of the Doppler effect, a siren will appear to rise in frequency as the vehicle travels towards you and fall as it moves away.

PHYSICSFILE

The sound of the Doppler effect

This Doppler effect behaviour can be easily modelled. You should be able to mimic the sound of a high-powered racing car like that in Figure 8.2.5 by making the sound 'neee…owwwww' with your voice. The 'neee' is the sound the racing car would make as it approached you—hence the high frequency. The 'owwwww' is the sound the racing car would make is it passed you and travelled away—hence the low frequency.

FIGURE 8.2.5 Daniel Ricciardo of Australia racing in Spain.

For a mechanical wave, the total Doppler effect may result from the motion of the source, the motion of the observer, or the motion of the medium the wave travels through. For waves that don't require a medium, such as light, only the relative difference in speed between the observer and the source will contribute to the effect.

EXTENSION

Doppler calculations

This study only requires a qualitative understanding of the Doppler effect as a wave phenomenon. However, as the relative motion between the observer and source is the cause of the change in apparent frequency, then by knowing what the relative motion is, the apparent frequency can be calculated.

In classical physics (which does not take into account relativistic effects), where the speed of the source and that of the observer are lower than the speed of the waves in the medium and the source and observer are approaching each other directly:

$$f = \left(\frac{v + v_0}{v - v_s}\right)f_0$$

where f is the apparent or observed frequency (Hz)

f_0 is the original frequency (Hz)

v is the speed of the waves in the medium (m s^{-1})

v_0 is the speed of the observer relative to the medium (m s^{-1}). v_0 is positive if the observer is moving towards the source and negative if moving away.

v_s is the speed of the source relative to the medium (m s^{-1}). v_s is positive if the source is moving towards the observer and negative if moving away.

As an approximation, if the speeds of the source and observer are small relative to the speed of the wave, then the approximate observed frequency is:

$$f = \left(1 + \frac{\Delta v}{v}\right)f_0, \text{ where } \Delta v = v_0 - v_s$$

and the approximate apparent change in frequency is

$$\Delta f = \left(\frac{\Delta v}{v}\right)f_0, \text{ where } \Delta f = f - f_0$$

An interesting additional effect was predicted by Lord Rayleigh. He predicted that if the source is moving at double the speed of sound, a musical piece emitted by the source would be heard in correct time and frequency, but *backwards*. Try to establish whether his prediction is true mathematically using the formulas above.

8.2 Review

SUMMARY

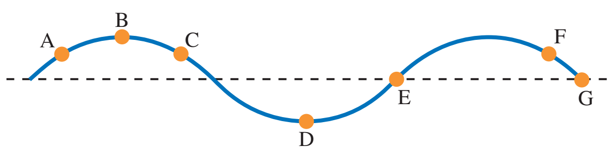
- Waves can be represented by displacement–distance graphs and displacement–time graphs.
- From a displacement–time graph, you can determine: amplitude, frequency and period.
- The period of a wave has an inverse relationship to the frequency, according to the relationship:

  $T = \frac{1}{f}$
- The speed of a wave can be calculated using the wave equation:

  $v = f\lambda = \frac{\lambda}{T}$
- The Doppler effect is a phenomenon that is observed whenever there is relative movement between the source of waves and an observer. It causes an apparent *increase* in frequency when the relative movement is *towards* the observer and an apparent *decrease* in frequency when the relative movement is *away* from the observer.
- For a mechanical wave, the total Doppler effect may result from the motion of the source, the motion of the observer, or the motion of the medium.

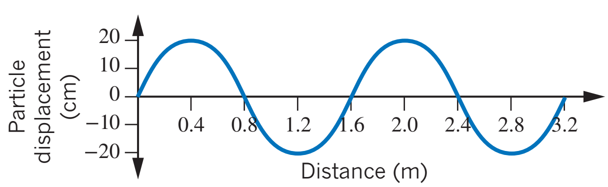
KEY QUESTIONS

1 Using the displacement–distance graph below, give the correct word or letters for the following:

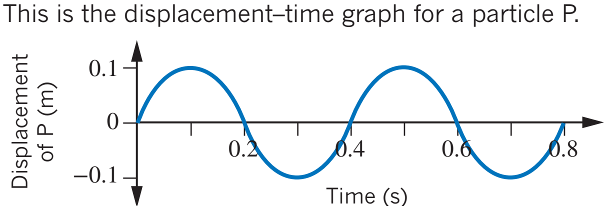


a two points on the wave that are in phase
b the name for the distance between these two points
c the two particles with maximum displacement from their rest position
d the term for this maximum displacement.

2 Use the graph below to determine the wavelength and the amplitude of this wave.



3 This is the displacement–time graph for a particle P.

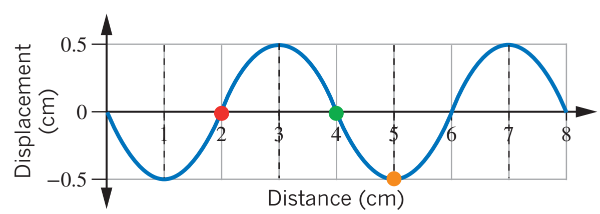


What is the:
a period of the wave
b frequency of the wave?

4 Five wavelengths of a wave pass a point each second. The amplitude is 0.3 m and the distance between successive crests of the waves is 1.3 m. What is the speed of the wave?

5 Which of the following is true and which is false? For the false statements rewrite them to make them true.
a The frequency of a wave is inversely proportional to its wavelength.
b The period of a wave is inversely proportional to its wavelength.
c The amplitude of a wave is not related to its speed.
d Only the wavelength of a wave determines its speed.

6 Consider the displacement–distance graph below.



a State the wavelength and amplitude of the wave.
b If the wave moves through one wavelength in 2 s, what is the speed of the wave?
c If the wave is moving to the right, which of the coloured particles is moving down?

7 Calculate the period of a wave with frequency 2×10^5 Hz.

8 A police car, travelling at 100 km h^{-1} along a straight road, has its siren sounding. The police car is pursuing another car travelling in the same direction, also at 100 km h^{-1}. There is no wind at the time. Would an observer in the car being pursued hear the siren from the police car at a higher, lower or the same frequency as it emits? Explain your answer.

9 An ambulance sounding its siren in still air moves towards you, then passes you and continues to move away in a straight path. How would the siren sound to you?

8.3 Wave interactions

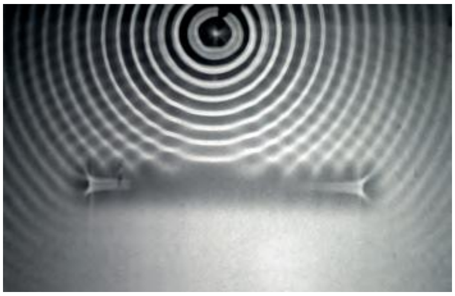


FIGURE 8.3.1 The reflection of waves in a ripple tank when meeting a solid surface, in this case a barrier positioned below the source of the circular waves.

Mechanical waves transfer energy through a medium. A medium is necessary, but there will be times when that medium physically ends, such as when a water wave meets the edge of a pool or air meets a wall. A change in the physical characteristics in the same medium, such as density and temperature, can act like a change in medium. When the medium ends, or changes, the wave doesn't just stop. Instead, the energy that the wave is carrying will undergo three processes:

- some energy will be *reflected* (see Figure 8.3.1)
- some energy will be *absorbed* by the new medium
- some energy will be *transmitted*.

Rarely in the real world does one mechanical wave occur in isolation. From the ripples that form on a pond hit by raindrops to the complex interactions of multiple reflected sound waves, the world is full of mechanical waves. The sounds produced by acoustic musical instruments and the human voice come from the interaction between sound waves and their reflections. The interaction of mechanical waves results in **superposition** and creates, among many other things, the characteristic sounds of musical instruments and of the human voice.

REFLECTION

When a transverse wave pulse reaches a hard surface, such as the fixed end of a rope, the wave is bounced back or **reflected**.

When the end of the rope is fixed, the reflected pulse is inverted (see Figure 8.3.2). So, for example, a wave crest would be reflected as a trough.

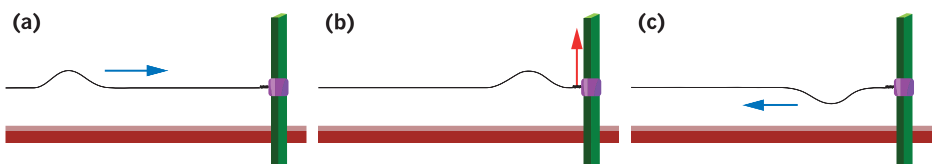


FIGURE 8.3.2 (a) A wave pulse moves along a string to the right and approaches a fixed post. (b) On reaching the end, the string exerts an upwards force on the fixed post. Due to Newton's third law, the fixed post exerts an equal and opposite force on the string which (c) inverts the wave pulse and sends its reflection back to the left on the bottom side of the string. There is a phase reversal on reflection from a fixed end.

This inversion can also be referred to as a 180° change of phase or, expressed in terms of the wavelength, λ, a shift in phase of $\frac{\lambda}{2}$.

When a wave pulse hits the end of the rope that is free to move (known as a free boundary), the pulse returns with no change of phase (see Figure 8.3.3). That is, the reflected pulse is the same as the incident pulse. A crest is reflected as a crest and a trough is reflected as a trough.

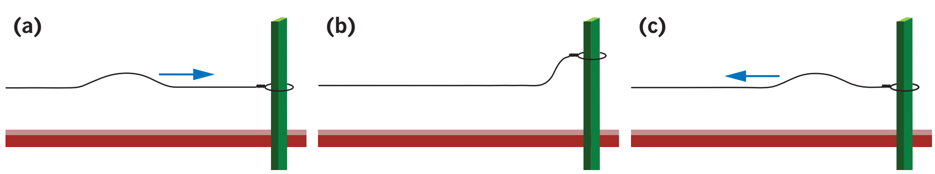


FIGURE 8.3.3 (a) A wave pulse moves along a string to the right and approaches a free end at the post. (b) On reaching the post the free end of the string is free to slide up the post. (c) No inversion happens and the wave pulse is reflected back to the left on the same side of the string, i.e. there is no phase reversal on reflection from a free end.

When the transverse wave pulse is reflected, the amplitude of the reflected wave isn't quite the same as the original. Part of the energy of the wave is **absorbed** by the post, where some will be transformed into heat energy and some will continue to travel through the post. You can see this more clearly by connecting a heavier rope to a lighter rope. The change in density has the same effect as a change in medium (see Figure 8.3.4).

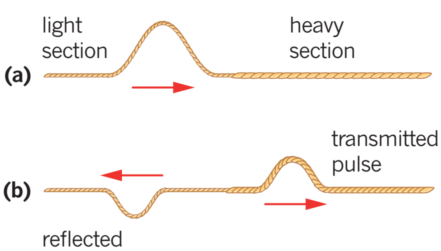


FIGURE 8.3.4 (a) A wave pulse travels along a light rope towards a heavier rope. (b) On reaching a change in density the wave pulse will be partly reflected and partly transmitted. This is analogous to a change in medium.

When a transverse wave pulse is sent down the rope from the light rope to the heavier rope, part of the wave pulse will be reflected and part of it will be **transmitted** to the heavier rope. As the second rope is heavier, a smaller proportion of the wave is transmitted into it and a larger proportion of the wave is reflected back.

This is just the same as a wave pulse striking a wall. The more rigid and/or dense the wall, the more the wave energy will be reflected and the less it will be absorbed – but there will always be some energy that is absorbed by or transferred to the second medium. This explains why sound can travel through walls.

Reflected wave fronts

Two- and three-dimensional waves, such as water waves, travel as **wave fronts**. When drawing wave fronts (see Figure 8.3.5), it is common to show the crests of the waves. When close to the source, wave fronts can show considerable curvature or may even be spherical when generated in three dimensions. Where a wave has travelled a long distance from its source, the wave front is nearly straight and is called a **plane wave**. A plane wave is shown in Figure 8.3.5(b). Plane waves can also be generated by a long, flat source such as those often used in a ripple tank.

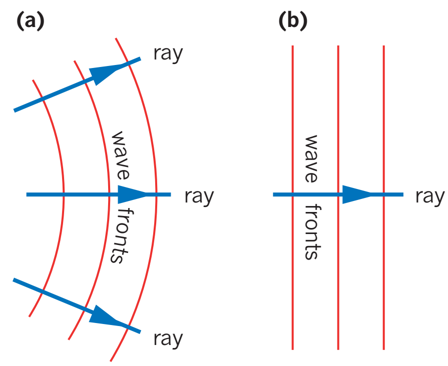


FIGURE 8.3.5 Rays can be used to illustrate the direction of motion of a wave. They are drawn perpendicular to the wave front of a two- or three-dimensional wave and in the direction of travel of the wave; (a) illustrates rays for circular waves near a point source while (b) shows a ray for plane waves.

The direction of motion of any wave front can be represented by a line drawn perpendicular to the wave front and in the direction the wave is moving (see Figure 8.3.5(a). This is called a **ray**. Rays can be used to study or illustrate the properties of two- and three-dimensional waves without the need to draw individual wave fronts.

By using rays to illustrate the path of a wave front reflecting from a surface, it can be shown that for a two- or three-dimensional wave, the angle from the **normal** at which the wave strikes a surface will equal the angle from the normal to the reflected wave. The normal is an imaginary line at 90°, i.e. perpendicular, to the surface.

These angles of the incident and reflected waves from the normal are labelled θ_i and θ_r, respectively, in Figure 8.3.6.

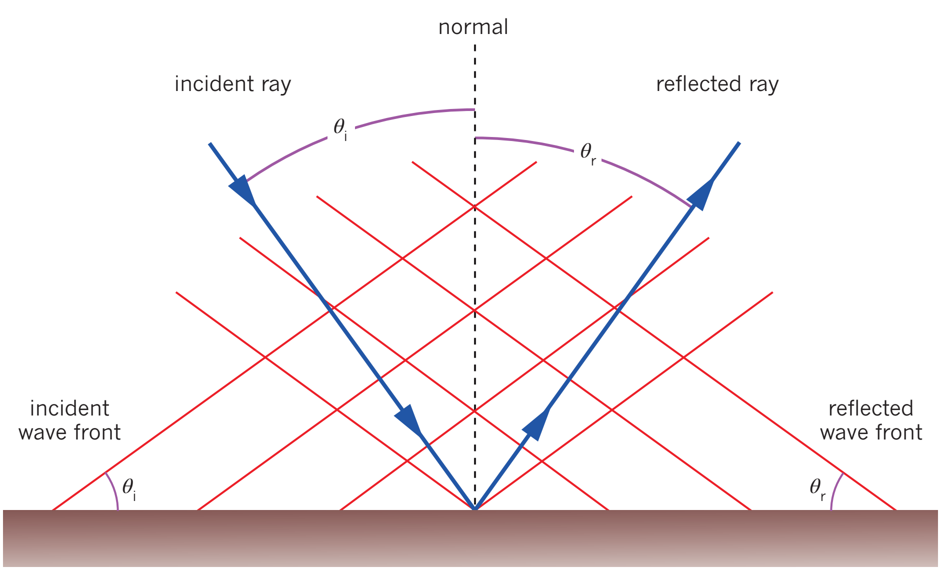


FIGURE 8.3.6 The law of reflection. The angle between the incident ray and the normal (θ_i) is the same as the angle between the normal and the reflected ray (θ_r).

This is referred to as the law of reflection. The law of reflection states that the **angle of reflection**, measured from the normal, equals the **angle of incidence** measured from the normal; that is $\theta_i = \theta_r$.

The law of reflection is true for any surface whether it is straight, curved or irregular. For all surfaces, including curved or irregular surfaces, the normal is drawn perpendicular to the surface at the point of contact of the incident ray or rays.

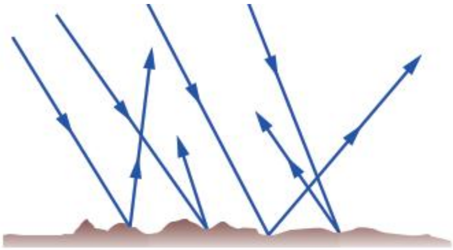


FIGURE 8.3.7 Reflection from an irregular surface. Each incident ray may be reflected in a different direction, depending upon how rough or irregular the reflecting surface is. The resulting wave will be diffuse (spread out).

When wave fronts meet an irregular, rough surface, the resulting reflection can be spread over a broad area. This is because each point on the surface may reflect the portion of the wave front reaching it in a different direction, as seen in Figure 8.3.7. This is referred to as **diffuse** reflection.

Echoes provide the most obvious evidence that sound waves are reflected. Like all waves, sound can be reflected when it strikes an obstacle.

PHYSICS IN ACTION

Sonic Depth Finder

The phenomenon of an echo is sometimes put to good use—an echo-sounding device makes it possible to measure the depth of the sea (Figure 8.3.8). This equipment works by measuring the time taken for the echo, generated when a sound wave is reflected by the sea bed, to return to the ship.

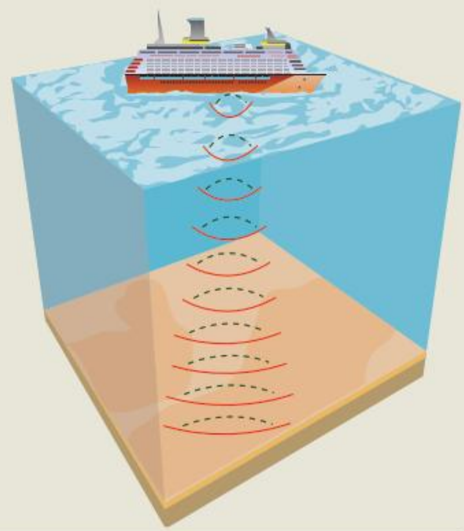


FIGURE 8.3.8 The ship sends a sound wave into the water below the ship. By measuring the time between the emitted sound (red waves) and its reflection (dashed lines), the ship's sonar can determine the depth of the water under the ship.

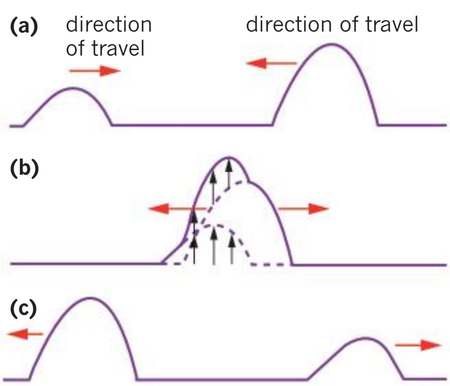


FIGURE 8.3.9 (a) As two wave pulses approach each other superposition occurs. (b) Shows the occurrence of constructive superposition. (c) After the interaction, the pulses continue unaltered; they do not permanently affect each other.

SUPERPOSITION

Imagine two transverse mechanical waves travelling towards each other along a string, as shown in Figure 8.3.9(a). When the crest of one wave coincides with the crest of the other, the resulting displacement of the string is the vector sum of the two individual displacements (see Figure 8.3.9(b)). The amplitude at this point is increased and the shape of the string resembles a combination of the two pulses. After they interact, the two pulses continue unaltered (see Figure 8.3.9(c)). The resulting pattern is a consequence of the principle of superposition. In this case, as the two waves added together, constructive superposition occurred.

When a pulse with a positive displacement meets one with a negative displacement as shown in Figure 8.3.10, the resulting wave will have a smaller amplitude (see Figure 8.3.10(b)). Once again the resulting displacement of the string is the vector sum of the two individual displacements; in this case a negative displacement adds to a positive displacement to produce a wave of smaller magnitude. This is called destructive superposition. Once again, the pulses emerge from the interaction unaltered (see Figure 8.3.10(c)).

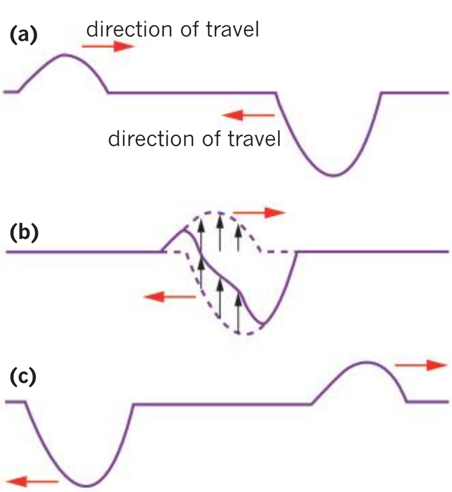


FIGURE 8.3.10 (a) As two wave pulses approach each other superposition occurs. (b) Superposition of waves in a string showing destructive superposition. (c) As in constructive superposition, the waves do not permanently affect each other.

When two waves meet and combine, there will be places where constructive superposition occurs and places where destructive superposition occurs. Although the wave pulses interact when they meet, passing through each other does not permanently alter the shape, amplitude or speed of either pulse. Just like transverse waves, longitudinal waves will also be superimposed as they interact.

The effects of superposition can be seen in many everyday examples. The ripples in the pond in Figure 8.3.11 were caused by raindrops hitting a pond. Where two ripples meet, a complex wave results from the superposition of the two waves, after which the ripples continue unaltered. Similarly in a crowded room, all the sounds reaching your ear are superimposed, so that one complex sound wave arrives at the eardrum.

Superposition is important both theoretically and practically in the formation of complex sounds. Imagine two single-frequency sound waves, or pure tones, one of which is twice the frequency of the other. The two individual waves are added together to give a more complicated resultant sound wave, as shown in Figure 8.3.12. Where one sound wave has a greater amplitude, as in the example illustrated, it will be the predominant sound heard. The quieter, higher frequency sound will combine with the louder one to create the sound that we hear. Note that for Figure 8.3.12, a transverse wave is used to depict the sound wave. The crests represent compressions (areas of high pressure) and the troughs represent rarefactions (areas of low pressure). Showing longitudinal waves in this way makes features, such as period and therefore frequency, easier to see.

FIGURE 8.3.11 The ripples from raindrops striking the surface of a pond behave independently regardless of whether they cross each other or not. Where the ripples meet, a complex wave will be seen as the result of the superposition of the component waves. After interacting, the component waves continue unaltered.

PHYSICSFILE

The cocktail party effect

In a crowded room, individual sound waves will interfere with each other repeatedly, but it is still possible to distinguish which person is speaking. If you know the person's voice, then you know that their voice will sound the same. To discern one person's speech amid all the sounds in the room, your brain uses an innate ability to 'undo' the superposition of waves by selecting one person's voice and suppressing all the other noise. The 'cocktail party effect' also highlights the ability to hear your own name over the noise of a group of people talking.

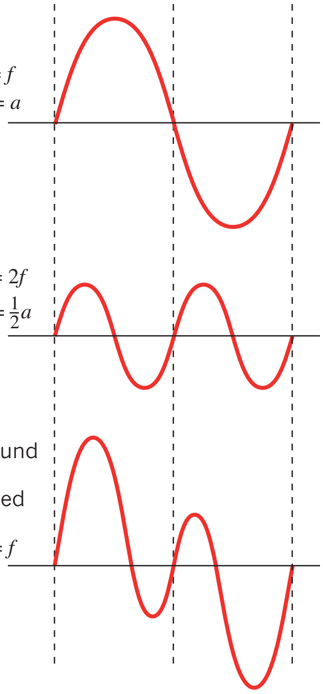


FIGURE 8.3.12 Two sound waves, one twice the frequency of the other, produce a complex wave of varying amplitude when they are superimposed.

RESONANCE

You may have heard about singers who supposedly can break glass by singing particularly high notes. Figure 8.3.13 shows a glass being broken in much the same way. All objects that can vibrate tend to do so at a specific frequency known as their natural or resonant frequency. **Resonance** is when an object is exposed to vibrations at a frequency equal to their resonant frequency. Resonance occurs when a weak vibration from one object causes a strong vibration in another. If the amplitude of the vibrations becomes too great, the object can be destroyed.

FIGURE 8.3.13 A glass can be destroyed by the vibrations caused by a singer emitting a sound of the same frequency as the resonant frequency of the glass.

A swing pushed once and left to swing or oscillate freely is an example of an object vibrating at its natural frequency. The frequency at which it moves backwards and forwards depends entirely on the design of the swing, mostly on the length of its supporting ropes. In time, the oscillations will fade away as the energy is transferred to the supporting frame and the air.

FIGURE 8.3.14 The sound box of a stringed instrument is tuned to resonate for the range of frequencies of the vibrations being produced by the strings. When a string is plucked or bowed, the airspace inside the box vibrates in resonance with the natural frequency and the sound is amplified.

If you watch a swing in motion, it is possible to determine its natural oscillating frequency. It is then possible to push the swing at exactly the right time so that you match its natural oscillation. The additional energy you add by pushing will increase the amplitude of the swing rather than work against it. Over time, the amplitude will increase and the swing will go higher and higher; this is resonance. The swing can only be pushed at one particular rate to get this increase in amplitude (i.e. to get the swing to resonate). If the rate is faster or slower, the forcing frequency that you are providing will not match the natural frequency of the swing and you will be fighting against the swing rather than assisting it.

Other examples of resonant frequency that you may have encountered are blowing air across the mouthpiece of a flute or drawing a bow across a string of a violin in just the right place (Figure 8.3.14). In each case, a clearly amplified sound is heard when the frequency of the forcing vibration matches a natural resonant frequency of the instrument.

Two very significant effects occur when the natural resonant frequency of an object is matched by the forcing frequency.

- The amplitude of the oscillations within the resonating object will increase dramatically.
- The maximum possible energy from the source creating the forced vibration is transferred to the resonating object.

PHYSICSFILE

Tacoma Narrows Gorge suspension bridge

Resonance was responsible for destroying a suspension bridge over the Tacoma Narrows Gorge in the US State of Washington in 1940. Wind gusts of 70 km h^{-1} caused vibrations with a forcing frequency that caused the bridge to oscillate with ever-increasing amplitude, until the whole bridge shook itself apart. That is, the gusts of wind provided a forcing frequency that perfectly matched the natural oscillating frequency of the bridge. This caused the bridge to vibrate more and more until eventually it was destroyed.
You can find video clips of the bridge falling if you search for it online.

In musical instruments and loudspeakers, resonance is a desired effect. The sounding boards of pianos and the enclosures of loudspeakers are designed to enhance and amplify particular frequencies. In other systems, such as car exhaust systems and suspension bridges, resonance is not always desirable, and care is taken to design a system that prevents resonance.

8.3 Review

SUMMARY

- A wave reaching the boundary between two materials in which it can travel will always be partly reflected, partly transmitted and partly absorbed.
- A wave has been reflected if it bounces back after reaching a boundary or surface.
- Waves reflect with a 180° phase change from fixed boundaries. That is, crests reflect as troughs and troughs reflect as crests.
- Waves reflect with no phase change from free boundaries. That is, crests reflect as crests and troughs as troughs.
- When a wave is reflected from a surface, the angle of reflection will equal the angle of incidence.
- The principle of superposition states that when two or more waves interact, the resultant displacement or pressure at each point along the wave will be the vector sum of the displacements or pressures of the component waves.
- Resonance occurs when the frequency of a forcing vibration equals the natural frequency of an object.
- Two special effects occur with resonance:
  - the amplitude of vibration increases
  - the maximum possible energy from the source is transferred to the resonating object.

KEY QUESTIONS

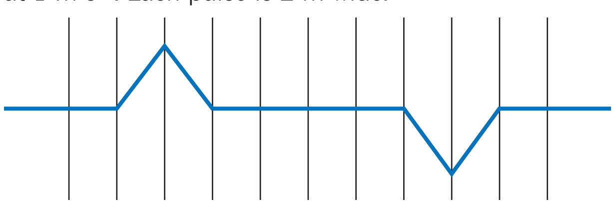
1 A wave travels along a rope and reaches the fixed end of the rope. What occurs next?

2 Which of the following properties of a wave can change when the wave is reflected: frequency, amplitude, wavelength or speed?

3 Which of the following about wave pulses are true and which are false? For the false statements, rewrite them so they are true.

- **a** The displacement of the resultant pulse is equal to the sum of the displacements of the individual pulses.
- **b** As the pulses pass through each other, the interaction permanently alters the characteristics of each pulse.
- **c** After the pulses have passed through each other, they will have the same characteristics as before the interaction.

4 Two triangular wave pulses head towards each other at 1 m s^{-1}. Each pulse is 2 m wide.

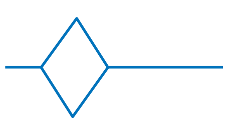


What will the superposition of these two pulses look like in 3 s?

A

B

C



D

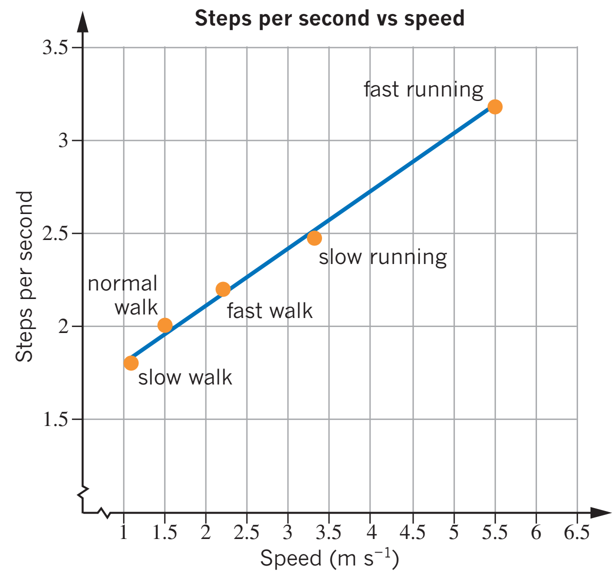
5 Explain why resonance can result in damage to man-made structures.

6 A light ray strikes a flat surface at an angle of 38° measured from the surface. What is the angle of reflection of the ray?

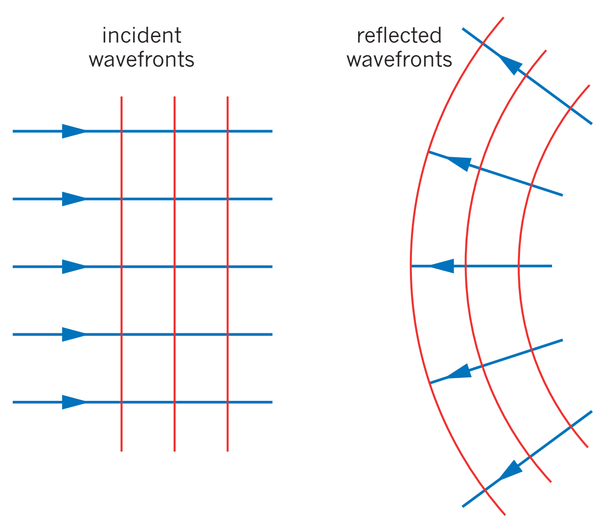
7 Resonance occurs when the frequency of a forcing vibration exactly equals the natural frequency of vibration of an object. Two special effects occur. Which one of the following responses relating to the effects of resonance is true? Explain your answer.

A The amplitude of vibration will decrease.
B The amplitude of vibration will increase.
C The frequency of vibration will increase.
D The frequency of vibration will decrease.

8 A footbridge over a river has a natural frequency of oscillation from side to side of approximately 1 Hz. When pedestrians walk at a pace that will produce an oscillation in the bridge close to the natural frequency of the footbridge, resonance will occur. The graph below displays relevant data about pedestrians walking or running. A pedestrian completes 1 cycle of their motion every 2 steps. Which activity of the pedestrians is most likely to cause damage to the footbridge over time? Explain your answer.



9 The following diagram shows a wave before and after being reflected from an object.



What is the shape of the object?

A flat
B concave
C convex
D parabolic

8.4 Standing waves in strings

FIGURE 8.4.1 Transverse standing waves can form along a violin string when the string is bowed by the violinist.

Drawing a bow across a violin string causes the string to vibrate between the fixed bridge of the violin and the finger of the violinist (see Figure 8.4.1). The simplest vibration will have maximum amplitude at the centre of the string, halfway between bridge and finger. This is a very simple example of a transverse **standing wave**.

Standing waves are formed from the superposition of waves. They occur when two waves of the same amplitude and frequency are travelling in opposite directions towards each other in the same string. Usually, one wave is the reflection of the other. Standing waves are responsible for the wide variety of sounds associated with speech and music.

STANDING WAVES IN A STRING

In Section 8.3 it was shown that when a wave pulse reaches a fixed end, it is reflected back 180° out of phase. That is, crests are reflected as troughs and troughs are reflected as crests.

Imagine creating a series of waves in a rope by shaking it vigorously. As the rope continues to be shaken, waves will travel in both directions. The new waves travelling down the rope will interfere with those being reflected back along the rope. This kind of motion will usually create quite a random pattern with the waves quickly dying away. Shaking the rope at just the right frequency, however, will create a new wave that interferes with the reflection in such a way that the two superimposed waves create a single, larger amplitude standing wave.

It is called a standing wave because the wave doesn't appear to be travelling along the rope. The rope simply seems to oscillate up and down with a fixed pattern. That is, is seems to be just 'flipping' up and down in a fixed pattern. This situation contrasts with a standard transverse wave where every point on the rope would have a maximum displacement at some time as the wave travels along the rope.

In Figure 8.4.2(a)–(d), two waves (drawn in blue) are shown travelling in opposite directions towards each other along a rope. One of the waves is a string of pulses (shown as a solid line) and the other is its reflection (shown as a dashed line). The two waves superimpose when they meet. Since the amplitude and frequency of each is the same, the end result, shown in part (e), is a standing wave. At the points where destructive interference occurs, the two waves totally cancel each other out and the rope will remain still. These are called **nodes**. Where the rope oscillates with maximum amplitude, constructive interference is occurring. These points on the standing wave are called **antinodes**.

Nodes and antinodes in a standing wave remain in a fixed position for a particular frequency of vibration. Figure 8.4.3 illustrates a series of possible standing waves in a rope, with both ends fixed, corresponding to three different frequencies. The lowest frequency of vibration (a) produces a standing wave with one antinode in the centre of the rope. The ends are fixed so they will always be nodal points. Assuming the tension in the rope is kept the same, patterns (b) and (c) are produced at twice and three times the original frequency respectively.

The rope could also vibrate at a frequency four times that of the original and so on. The frequencies at which standing waves are produced are called the resonant frequencies of the rope.

It is important to note that the formation of a standing wave does not mean that the string or rope itself is stationary. It will continue to oscillate as further wave pulses travel up and down the rope. It is the relative position of the nodes and antinodes that remain unchanged.

It is also important to note that standing waves are not a natural consequence of every wave reflection.

Standing waves are only produced by the superposition of two waves of equal amplitude and frequency, travelling in opposite directions.

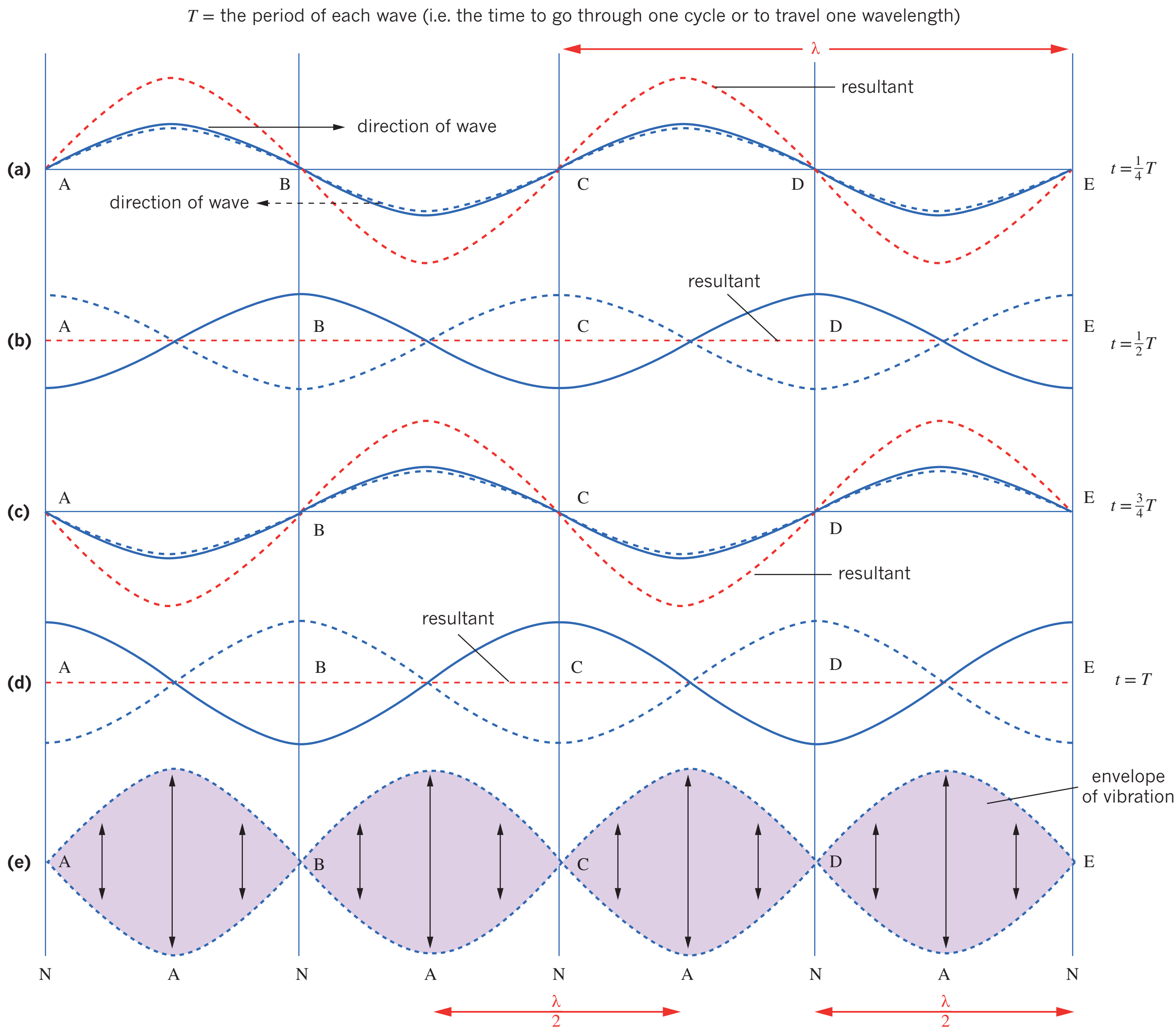


FIGURE 8.4.2 A standing wave created in a rope from two waves travelling in opposite directions, each with the same amplitude and frequency. (a) At a particular point in time, the two waves are completely superimposed, resulting in a wave twice the original amplitude. (b) After a time equal to $\frac{T}{4}$ (one-quarter of the period), the waves will each have moved $\frac{\lambda}{4}$, which means that they have moved $\frac{\lambda}{2}$ in relation to each other. The waves are completely out of phase and the resulting displacement is zero. (c) and (d) As more time goes by, the waves will continue to move past each other and completely superimpose again before cancelling again. (e) The cycles shown in (a) to (d) form a standing wave. A standing wave swings between maximum displacements, creating antinodes (A) which lie halfway between the stationary nodes (N). Regardless of the position of the component waves, these nodes stay in the same place. Successive nodal points lie $\frac{\lambda}{2}$ apart, as do successive antinodal points.

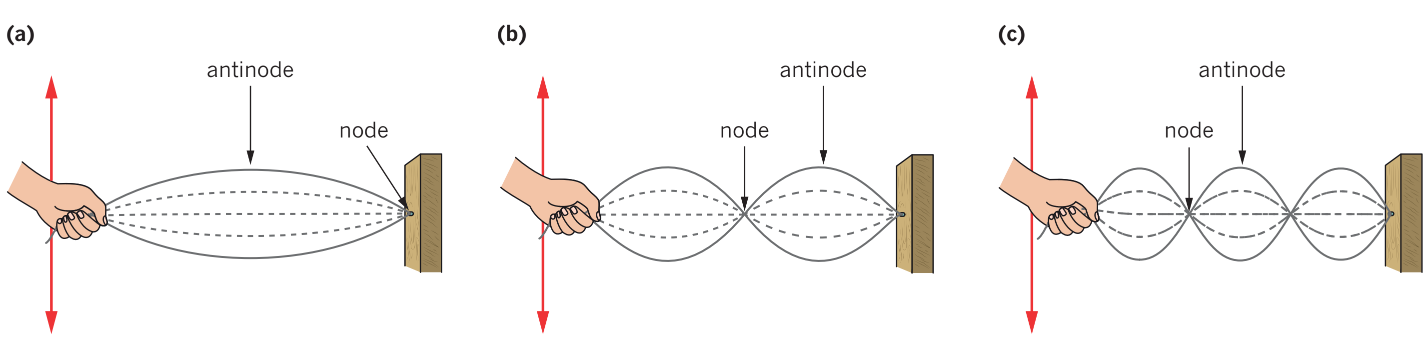


FIGURE 8.4.3 A rope vibrated at three different resonant frequencies, illustrating the standing waves produced at each frequency.

Standing waves are a result of resonance and occur only at the natural frequencies of vibration, or resonant frequencies, of the particular medium.

Standing waves don't only exist in the everyday world but also occur at the sub-atomic level. This concept is covered in detail in Chapter 10.

HARMONICS

From the strings of a musical instrument, a large variety of waves are created. They travel along the string in both directions and reflect from the fixed ends. Most of these vibrations will interfere in a random fashion and die away. However, those corresponding to resonant frequencies of the string will form standing waves and remain.

The resonant frequencies produced in this complex vibration of multiple standing waves are termed **harmonics**. The lowest and simplest form of vibration, with one antinode (see Figure 8.4.3(a)), is called the **fundamental** frequency. Higher-level harmonics (see Figure 8.4.3(b) and (c)) are referred to by musicians as **overtones**.

The fundamental frequency usually has the greatest amplitude, so it has the greatest influence on the sound. The amplitude generally decreases for each subsequent harmonic. Usually all possible harmonics are produced in a string simultaneously, and the instrument and the air around it also vibrate to create the complex mixture of frequencies heard as an instrumental note.

The resonant frequencies or harmonics in a string of length l can be calculated from the relationship between the length of the string and the wavelength, λ, of the corresponding standing wave.

For a string fixed at both ends:

The first harmonic, or fundamental frequency, has one antinode in the centre of the string and

$$\lambda = 2 \times l$$

The second harmonic will have two antinodes and

$$\lambda = l = \frac{2l}{2}$$

The third harmonic will have three antinodes

$$\lambda = \frac{2l}{3}$$

And so in general, for any harmonic:

$$\lambda = \frac{2l}{n}$$

where λ is the wavelength (m)

l is the length of the string (m)

n is the number of the harmonic, which is also the number of antinodes (i.e. 1, 2, 3, 4...)

The relationship between wavelength, λ, and string length, l, is shown in Figure 8.4.4.

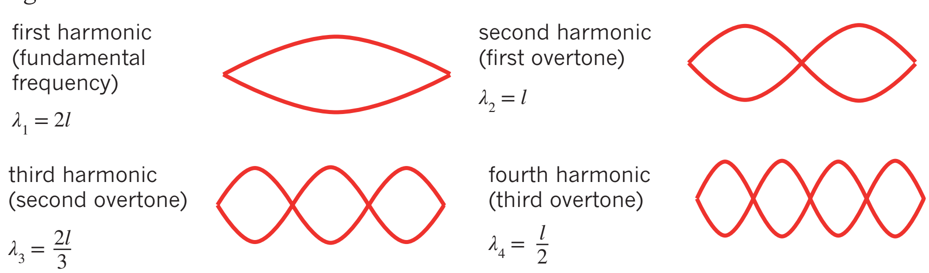


FIGURE 8.4.4 The first four resonant frequencies, or harmonics, in a stretched string fixed at both ends. The ends are fixed so they will always be nodal points.

Using the wave equation $v = f\lambda$ gives the relationship between frequency, velocity and string length.

For the first harmonic, or fundamental frequency:

$$v = f\lambda \text{ and so } f = \frac{v}{\lambda} = \frac{v}{2l}$$

For the second harmonic:

$$f = \frac{v}{\lambda} = \frac{v}{l}$$

For the third harmonic:

$$f = \frac{v}{\lambda} = \frac{3v}{2l}$$

And so in general:

$$f = \frac{nv}{2l}$$

where n is the number of the harmonic

f is the frequency of the wave (Hz)

v is the velocity of the wave (m s^{-1})

l is the length of the string (m)

For a string fixed at one end and free at the other, the standing waves that form are shown in Figure 8.4.5.

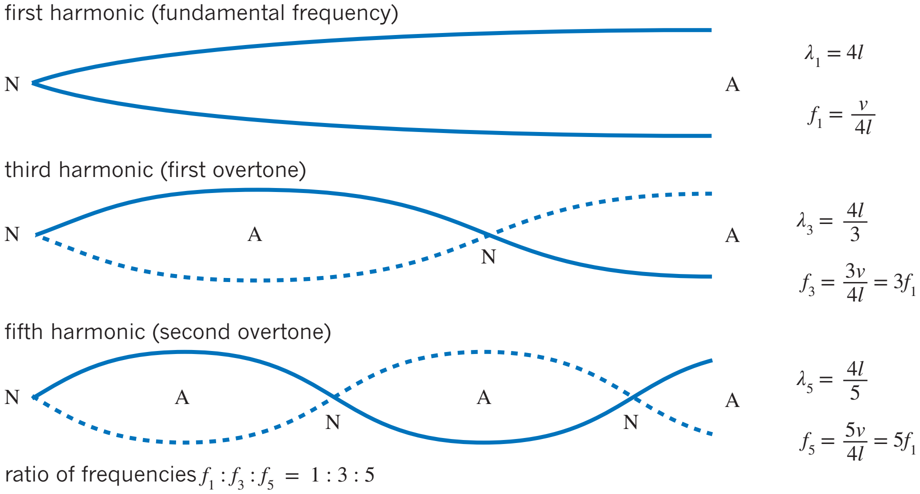


FIGURE 8.4.5 The lower harmonics for a string that is fixed at one end (the left-hand side of the diagram) and free to move at the other (the right-hand side of the diagram). Only odd-numbered harmonics are possible, since only these satisfy the condition of having a node at the fixed end and an antinode at the free end.

A node will always form at the fixed end (the left-hand side of Figure 8.4.5) and an antinode will always form at the free end (the right-hand side of Figure 8.4.5) of the string. The first harmonic, or fundamental frequency, will have a wavelength four times the length of the string:

$$\lambda = 4 \times l$$

The next simplest harmonic that can form will have a wavelength:

$$\lambda = \frac{4 \times l}{3}$$

The next harmonic that can form will have a wavelength:

$$\lambda = \frac{4 \times l}{5}$$

In general:

$$\lambda = \frac{4l}{n}$$

where λ is the wavelength (m)

l is the length of the string (m)

n is the number of the harmonic, odd-number integers only (i.e. 1, 3, 5…)

And in general, for frequency in terms of velocity:

$$f = \frac{nv}{4l}$$

where n is the number of the harmonic (i.e. 1, 3, 5…)

Notice that only odd-numbered harmonics will form since the conditions necessary for a standing wave to form are only met under the circumstances where there is an antinode at the free end and a node at the fixed end. The ratio of the wavelengths of the harmonics will be 1:3:5… That means the wavelength of the third harmonic is $\frac{1}{3}$ the length of the fundamental frequency (the first harmonic), the fifth harmonic is $\frac{1}{5}$ the length of the fundamental frequency, and so on.

Even numbered harmonics (i.e. 2, 4, 6…) will not form. The equations introduced above can be modified to incorporate this. The modified formula is:

$$\lambda = \frac{4l}{(2n-1)}$$

where λ is the wavelength in metres (m)

l is the length of the string in metres (m)

n is an integer (i.e. 1, 2, 3, 4...)

Note that for this equation, n is defined slightly differently than that for strings fixed at both ends. For this equation, n is the next harmonic in the sequence, not the harmonic number.

While it won't be considered in this unit, it should also be noted that the resonant frequencies of a string correspond to a particular tension and mass per unit length. Tightening or loosening the string will change the wavelengths and resonant frequencies for that string (i.e. the instrument will need tuning by adjusting the tension of the string). Heavier strings of a particular length will have different resonant frequencies than lighter strings of the same length and tension.

PHYSICSFILE

Surface waves

Seismic surface waves travel along the boundary between materials, such as the Earth's crust and upper mantle. One type of surface wave is called the Rayleigh wave, or ground roll. They are surface waves that travel as ripples with a motion like that of waves on the surface of water, although the restoring force is elastic rather than gravitational as it is for water waves. A phenomenon known as free oscillation of the Earth is the result of the superposition between two such surface waves travelling in opposite directions creating a surface standing wave.

The first observations of free oscillations of the Earth were made during the 1960 Chile earthquake. Since then thousands of harmonics have been identified.

Worked example 8.4.1

FUNDAMENTAL FREQUENCY

A violin string, fixed at both ends, has a length of 22 cm. It is vibrating at its fundamental mode of vibration at a frequency of 880 Hz.

a What is the wavelength of the fundamental frequency?	
Thinking	**Working**
Identify the length of the string (l) in metres and the harmonic number (n).	l = 22 cm = 0.22 m n = 1
Recall that for any frequency, $\lambda = \frac{2l}{n}$. Substitute the values from the question and solve for λ.	$\lambda = \frac{2l}{n}$ $= \frac{2 \times 0.22}{1}$ = 0.44 m

b What is the wavelength of the second harmonic?	
Thinking	**Working**
Identify the length of the string (l) in metres and the harmonic number (n).	l = 22 cm = 0.22 m n = 2
Recall that for any frequency $\lambda = \frac{2l}{n}$. Substitute the values from the question and solve for λ.	$\lambda = \frac{2l}{n}$ $= \frac{2 \times 0.22}{2}$ = 0.22 m

Worked example: Try yourself 8.4.1

FUNDAMENTAL FREQUENCY

A standing wave in a string is found to have a wavelength of 0.50 m for the fundamental frequency of vibration. Assume that the tension of the string is not changed and that the string is fixed at both ends.

a What is the length of the string?

b What is the wavelength of the third harmonic?

PHYSICS IN ACTION

Wind instruments, air columns and other standing waves

Longitudinal stationary waves are also possible in air columns. These create the sounds associated with wind instruments. Blowing over the hole of a flute (see Figure 8.4.6) or the reed of a saxophone produces vibrations that correspond to a range of frequencies that create sound waves in the tube.

The compressions and rarefactions of the sound waves, confined within the tube, reflect from both open and closed ends. This creates the right conditions for resonance and the formation of standing waves. The length of the pipe will determine the frequency of the sounds that will resonate.

The open end of a pipe corresponds to the fixed end of a string in that the reflected wave is fully inverted (i.e. it undergoes a 180° change of phase). At a closed end there is no change of phase. The reflected wave is in phase with the original wave.

FIGURE 8.4.6 Blowing over the mouthpiece of a flute and controlling the length of the flute with the keys enables a particular note to be produced.

While the discussion in this section has been of two-dimensional standing waves, standing waves may also form in three dimensions, such as in a section of Earth's crust. Standing waves will form as a result of resonance in any wave form.

8.4 Review

SUMMARY

- Standing or stationary waves occur as a result of resonance at the natural frequency of vibration.
- Points on a standing wave that remain still are called nodes.
- Points of maximum amplitude on a standing wave are called antinodes.
- The standing wave frequencies are referred to as harmonics. The simplest mode is referred to as the fundamental frequency.
- Within a string fixed at both ends, the wavelength of the standing waves corresponding to the various harmonics is:
  $\lambda = \frac{2l}{n}$
  and the frequency is:
  $f = \frac{nv}{2l}$
  All harmonics may be present.
- Within a string fixed at one end, the wavelength of the standing waves corresponding to the various harmonics is:
  $\lambda = \frac{4l}{n}$
  and the frequency is:
  $f = \frac{nv}{4l}$
  Only odd numbered harmonics may be present.

KEY QUESTIONS

1 A transverse standing wave is produced using a rope. Is the standing wave actually standing still? Explain your answer.

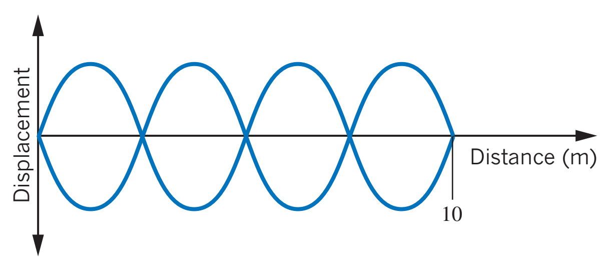
2 Describe how superposition and interference are related to the formation of standing waves in a stretched slinky spring.

3 What is the wavelength of the fundamental mode of a standing wave on a string 0.4 m long and fixed at both ends?

4 Calculate the length of a string fixed at both ends when the wavelength of the fourth harmonic is 0.75 m.

5 A standing wave is produced in a rope fixed at both ends by vibrating the rope with four times the frequency that produces the fundamental or first harmonic. How much larger or smaller is the wavelength of this standing wave compared to that of the fundamental or first harmonic?

6 A standing wave pattern in a string is shown over a distance of 10 m.



What is the length of the rope that would generate the first harmonic if a standing wave of the same wavelength shown in the diagram above was produced?

7 The fundamental frequency of a violin string is 350 Hz and the velocity of the waves along it is 387 m s^{-1}. What is the wavelength of the new fundamental when a finger is pressed to shorten the string to $\frac{2}{3}$ its original length?

The following information relates to questions 8–10.

A metal string (at constant tension) of length 50 cm is plucked, creating a wave pulse. The speed of the transverse wave created is 300 m s^{-1}. Both ends of the string are fixed.

8 Calculate the frequency of the fundamental frequency.

9 Calculate the frequency of the second harmonic.

10 Calculate the frequency of the third harmonic.

Chapter review

08

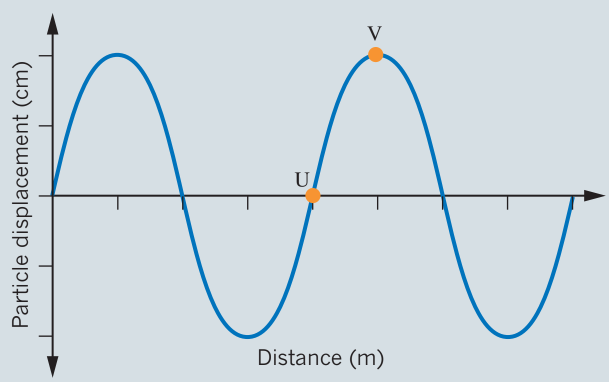
KEY TERMS

absorb
amplitude
angle of incidence
angle of reflection
antinode
compression
crest
diffuse
Doppler effect
frequency
fundamental
harmonic
longitudinal
mechanical wave
medium
node
normal
overtone
period
phase
plane wave
pulse
rarefaction
ray
reflect
resonance
sinusoidal
standing wave
superposition
transmit
transverse
trough
wave front
wavelength

1 Imagine that you watch from above as a stone is dropped into water. Describe the movement of the particles on the surface of the water.

2 Describe the similarities and differences between transverse and longitudinal waves.

3 At the moment in time shown on the graph, in what directions are the particles U and V moving?



4 The source of waves in a ripple tank vibrates at a frequency of 10.0 Hz. If the wave crests formed are 30.0 mm apart, what is the speed of the waves (in $m\ s^{-1}$) in the tank?

5 A submarine's sonar sends out a signal with a frequency of 32 kHz. If the wave travels at $1400\ m\ s^{-1}$ in seawater, what is the wavelength of the signal?

6 Assuming the speed of sound in water is $1500\ m\ s^{-1}$, what would be the wavelength of a sound of frequency 300 Hz?

7 The same sound wave from Question 6 is now produced in air where the speed of sound is $340\ m\ s^{-1}$. What is the wavelength now?

8 Two vehicles are driving in the same direction down a road at $60\ km\ h^{-1}$. The driver in the lead car sounds the horn, which is set at a frequency of 256 Hz. At what apparent frequency will the driver in the following car hear the sound?

9 A motor bike is able to produce a long, steady sound. You are unable to see the motor bike, but can hear the sound from it rise in frequency and then fall. Which one or more of the following options best explains the motion of the motor bike relative to you?

A The bike travelled towards you.
B The bike travelled away from you.
C The bike travelled past you.
D The bike travelled towards you, then away from you.

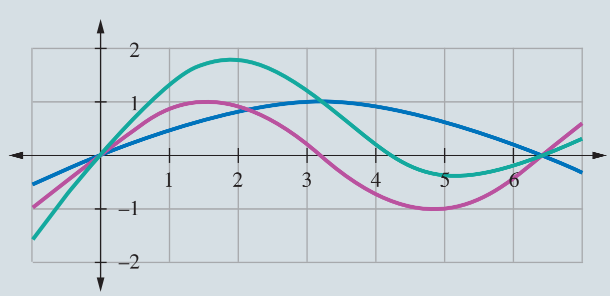
10 If you decreased the wavelength of the sound made by a loudspeaker, what effect would this have on the frequency and the velocity of the sound waves?

11 A pulse is sent along a rope that is fixed at one end. What property of the pulse changes when it is reflected at the fixed end of the rope?

12 When sunlight shines through a window in a house, its energy can be transmitted, reflected or absorbed. Which of these processes is responsible for the fact that:

a the interior of the house is illuminated
b when light falls on a window, you can see some of it from outside the house
c the glass gets warm?

13 The following graph shows three wave forms. Two of the wave forms superimpose to form the third wave form:



Which wave is the result of the superposition of the other two?

Chapter review *continued*

14 Using ideas about the movement of particles in air, explain how you know sound waves only carry energy and not matter from one place to another.

15 Describe the concept of resonance and why it would need to be considered when designing structures like buildings or bridges.

The following information relates to questions 16–18.

A signal generator is attached to a string vibrator producing vibrations down a string fixed at one end and free to move at the other. The string is kept at constant tension. The effective length of the string is 85 cm. The speed of the vibrations along the string is 340 m s^{-1}.

16 What is the lowest frequency of vibration that will produce a standing wave in the string?

17 What is the frequency of vibration of the third harmonic?

18 An earthquake causes a footbridge to oscillate up and down with a fundamental frequency once every 4.0 s. The motion of the footbridge can be considered to be like that of a string fixed at both ends. What is the frequency of the second harmonic for this footbridge?

19 The velocity of waves in a particular string at constant tension is 78 m s^{-1}. The string is fixed at both ends. If a particular frequency of a standing wave formed in the string is 428 Hz, how far apart would two adjacent antinodes be?

20 Reflection is possible from which of the following shaped surfaces if each surface is reflective? (More than one answer may be correct.)

A flat surface
B concave (curved in) surface
C uneven surface
D convex (curved out) surface

21 In a medium where a mechanical wave can be propagated, when is the Doppler effect observable?

CHAPTER 09

The nature of light

Discovering the nature of light has been one of the scientific community's greatest challenges. Over the course of history, light has been compared to a geometric ray, a stream of particles or even a series of waves. However, these relatively simple models have been found to be limited in their ability to explain all of the properties of light.

The search for a more adequate model has pushed scientists to develop new types of equipment and more-sophisticated experiments. Over time, it has also led to a reshaping of the understanding of the nature of matter and energy.

This chapter will follow the historical changes in the understanding of the nature of light, and will give a general overview of the properties of waves, Young's experiment and the other evidence that caused 19th-century scientists to develop a wave model for light.

Key knowledge

By the end of this chapter you will have studied the physics of the nature of light, and will be able to:

- describe light as an electromagnetic wave which is produced by the acceleration of charges, which in turn produces changing electric fields and associated changing magnetic fields
- identify that all electromagnetic waves travel at the same speed, c, in a vacuum
- compare the wavelength and frequencies of different regions of the electromagnetic spectrum, including radio, microwave, infrared, visible, ultraviolet, X-ray and gamma, and identify the distinct uses each has in society
- explain polarisation of visible light and its relation to a transverse wave model
- investigate and analyse theoretically and practically the behaviour of waves including:
  - refraction using Snell's law: $n_1 \sin \theta_1 = n_2 \sin \theta_2$ and $n_1 v_1 = n_2 v_2$
  - total internal reflection and critical angle including applications: $n_1 \sin \theta_c = n_2 \sin 90°$
- investigate and explain theoretically and practically colour dispersion in prisms and lenses with reference to refraction of the components of white light as they pass from one medium to another
- explain the results of Young's double-slit experiment with reference to:
  - evidence for the wave-like nature of light
  - constructive and destructive interference of coherent waves in terms of path differences: $n\lambda$ and $\left(n - \frac{1}{2}\right)\lambda$ respectively
  - effect of wavelength, distance of screen and slit separation on interference patterns: $\Delta x = \frac{\lambda L}{d}$
  - investigate and analyse theoretically and practically constructive and destructive interference from two sources with reference to coherent waves and path difference: $n\lambda$ and $\left(n - \frac{1}{2}\right)\lambda$ respectively
  - investigate and explain theoretically and practically diffraction as the directional spread of various frequencies with reference to different gap width or obstacle size, including the qualitative effect of changing the $\frac{\lambda}{w}$ ratio.

9.1 Light as a wave

One of the great scientific achievements of the 19th century was the development of a comprehensive wave model for light. This model was able to explain a large number of wave properties including reflection, refraction, dispersion (as shown in Figure 9.1.1), diffraction, interference and polarisation. This also led to a deeper understanding of phenomena such as heat and radio transmissions.

FIGURE 9.1.1 The wave model of light can explain the phenomenon of the dispersion of light into its component colours.

WAVE MODEL VERSUS PARTICLE MODEL

In the late 17th century a debate raged among scientists about the nature of light.

The famous English scientist Sir Isaac Newton explained light in terms of particles or 'corpuscles', with each different colour of the spectrum representing a different type of particle. Scientists Robert Hooke (from England, and who you might recall from Chapter 7 'The relationship between force, energy and mass') and Christiaan Huygens (from Holland) proposed an alternative model that described light as a type of wave, similar to the water waves observed in the ocean.

A key point of difference between the two theories was that Newton's 'corpuscular' theory suggested that light would speed up as it travelled through a solid material such as glass. In comparison, the wave theory predicted that light would be slower in glass than in air.

Unfortunately, at that time it was impossible to measure the speed of light accurately, so the question could not be resolved scientifically. Newton's esteemed reputation meant that for many years his corpuscular theory was considered correct.

It was not until the early 19th century that experiments first convincingly demonstrated the wave properties of light.

Today, a modern understanding of light draws on aspects of both theories and is, perhaps, more complex than either Newton, Hooke or Huygens could ever have imagined.

HUYGENS' PRINCIPLE

The theoretical basis for wave propagation in two dimensions was first explained by the Dutch scientist Christiaan Huygens. Huygens' principle states that each point on a wavefront can be considered as a source of secondary wavelets (i.e. small waves).

Consider the plane wave shown in Figure 9.1.2. Each point on the initial wavefront can be treated as if it is a point source producing circular waves, some of which are shown in green. After one period, these circular waves will have advanced by a distance equal to one wavelength. Huygens proved mathematically that when the amplitudes of each of the individual circular waves are added, the result is another plane wave as shown by the new wavefront.

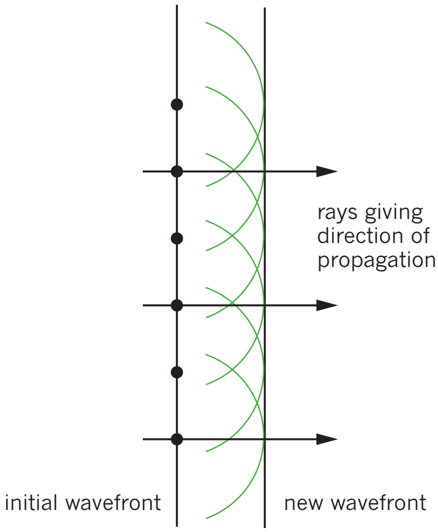


FIGURE 9.1.2 Each point on the wavefront of a plane wave can be considered as a source of secondary wavelets. These wavelets combine to produce a new plane wavefront.

This process is repeated at the new wavefront, causing the wave to propagate in the direction shown.

Circular waves are propagated in a similar way, as shown in Figure 9.1.3.

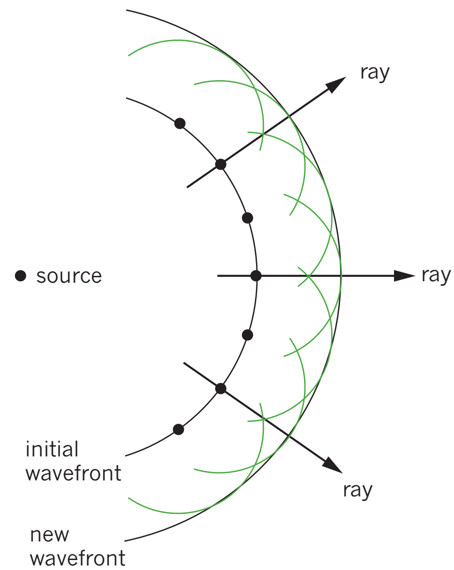
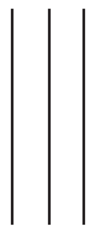


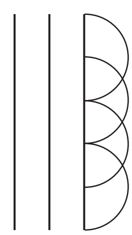
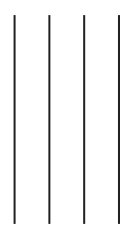
FIGURE 9.1.3 Each point on the wavefront of a circular wave can be considered as a source of secondary wavelets. These wavelets combine to produce a new circular wavefront.

Worked example 9.1.1

APPLYING HUYGENS' PRINCIPLE

On the plane wave shown moving from left to right below, sketch some of the secondary wavelets on the wavefront and draw the appearance of the new wave formed after one period.

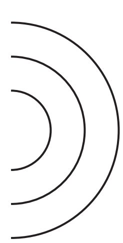


Thinking	Working
Sketch a number of secondary wavelets on the advancing wavefront.	
Sketch the new wavefront.	

Worked example: Try yourself 9.1.1

APPLYING HUYGENS' PRINCIPLE

On the circular waves shown below, sketch some of the secondary wavelets on the outer wavefront and draw the appearance of the new wave formed after one period.



A wave model can be used to explain a number of important properties of light including:

- refraction
- dispersion
- diffraction
- polarisation.

REFRACTION

Refraction is a change in the direction of light caused by changes in its speed. Changes in the speed of light occur when light passes from one medium (substance) into another. In Figure 9.1.4, the light changes direction as it leaves the glass prism and re-enters the air.

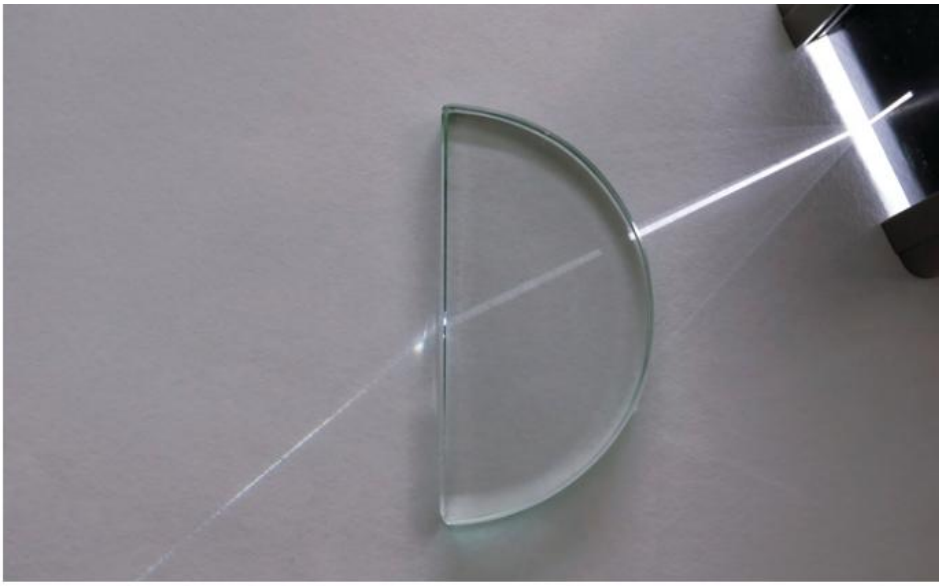


FIGURE 9.1.4 Light refracts as it moves from one medium (i.e. the semicircular glass prism) into another (i.e. air) causing a change in direction.

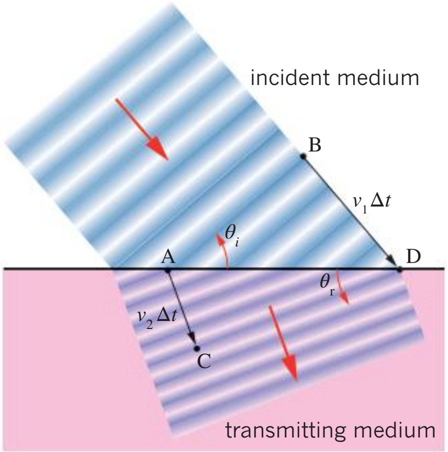


FIGURE 9.1.5 Wave refraction occurs because the distance A–C travelled by the wave in the transmitting medium is shorter than the distance B–D that it travels at the same time in the incident medium.

Consider Figure 9.1.5, where light waves are moving from an incident medium where they have high speed, v_1, into a transmitting medium where they have a lower speed, v_2. At the same time that the wave travels a distance $v_1\Delta t$ (B–D) in the incident medium, it travels a shorter distance $v_2\Delta t$ (A–C) in the transmitting medium. In order to do this, the wavefronts must change direction or 'refract' as shown.

Light waves behave in a similar way when they move from a medium like air into water. The direction of the refraction depends on whether the waves speed up or slow down when they move into the new medium. In Figure 9.1.6, the light waves slow down as they move from air into glass so the direction of propagation of the wave is refracted towards the normal. The angle of incidence, *i*, which is defined as the angle between the direction of propagation and the normal, is greater than the angle of refraction, *r*.

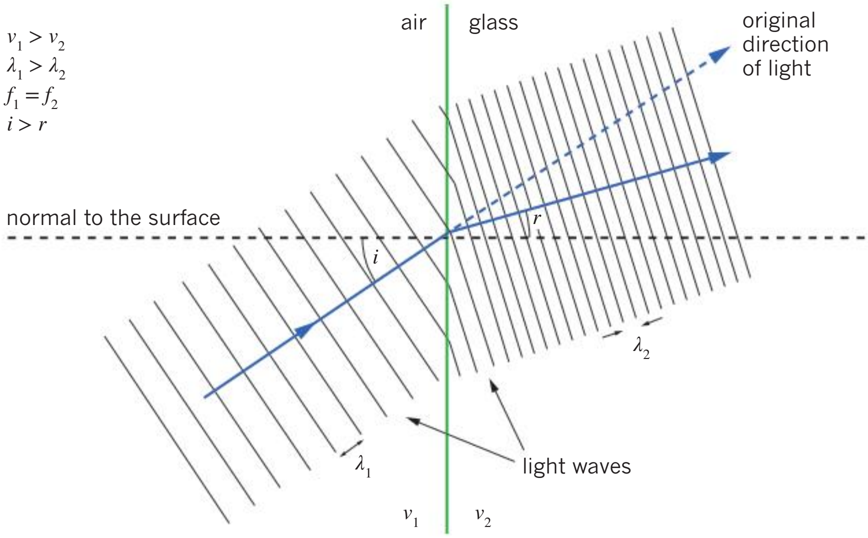


FIGURE 9.1.6 Light waves refract towards the normal when they slow down.

Conversely, when a light wave moves from glass where it has low speed into air where it travels more quickly, it is refracted away from the normal, as shown in Figure 9.1.7. In other words, the angle of incidence, i, is less than the angle of refraction, r.

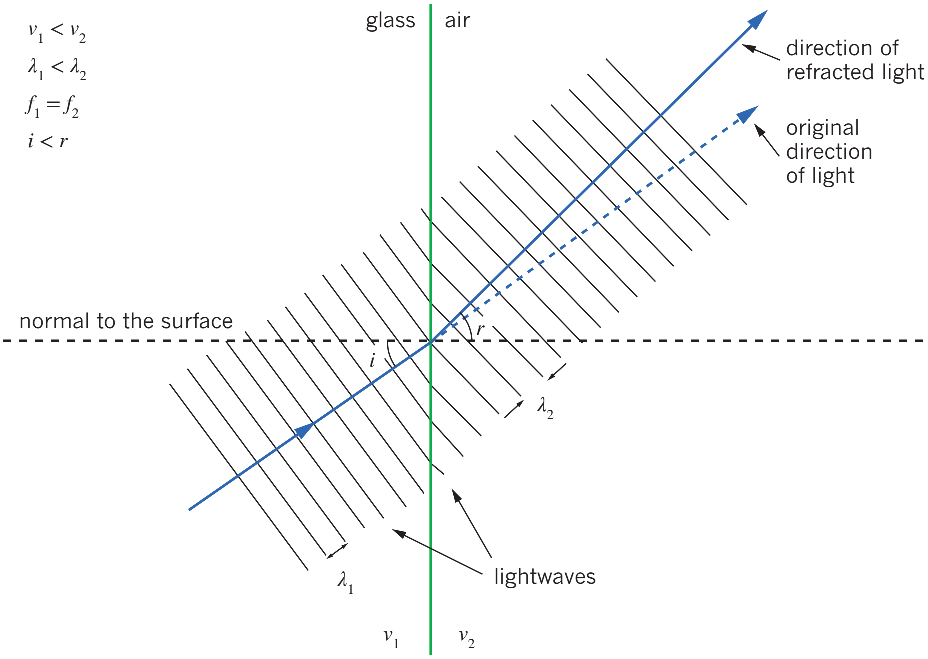


FIGURE 9.1.7 Light waves refract away from the normal when they speed up.

Note that when a wave changes its speed, its wavelength also changes correspondingly, but its frequency does not change as there are still the same number of waves; waves cannot be gained or lost.

Refractive index

The amount of refraction that occurs depends on how much the speed of light changes as light moves from one medium to another—when light slows down greatly, it will undergo significant refraction.

The speed of light in a number of different materials is shown in Table 9.1.1.

Material	Speed of light ($\times 10^8$ m s^{-1})
vacuum	3.00
air	3.00
ice	2.29
water	2.25
quartz	2.05
crown glass	1.97
flint glass	1.85
diamond	1.24

TABLE 9.1.1 The speed of light in various materials correct to three significant figures.

Scientists find it convenient to describe the change in speed of a wave using a property called the **refractive index**. The refractive index of a material, n, is defined as the ratio of the speed of light in a vacuum, c, to the speed of light in the medium, v:

$$n = \frac{c}{v}$$

Note that n is dimensionless, i.e. it has no units, it is just a number.

The refractive index for various materials is given in Table 9.1.2.

Material	Refractive index, *n*
vacuum	1.00
air	1.00
ice	1.31
water	1.33
quartz	1.46
crown glass	1.52
flint glass	1.62
diamond	2.42

TABLE 9.1.2 Refractive indices of various materials.

This quantity is also sometimes referred to as the 'absolute' refractive index of the material, to distinguish it from the 'relative' refractive index that might be used when a light ray moves from one medium to another, e.g. from water to glass.

Worked example 9.1.2

CALCULATING REFRACTIVE INDEX

The speed of light in water is 2.25×10^8 m s^{-1}. Given that the speed of light in a vacuum is 3.00×10^8 m s^{-1}, calculate the refractive index of water.

Thinking	Working
Recall the definition of refractive index.	$n = \frac{c}{v}$
Substitute the appropriate values into the formula and solve.	$n = \frac{3.00 \times 10^8}{2.25 \times 10^8} = \frac{3.00}{2.25} = 1.33$

Worked example: Try yourself 9.1.2

CALCULATING REFRACTIVE INDEX

The speed of light in crown glass (a type of glass used in optics) is 1.97×10^8 m s^{-1}. Given that the speed of light in a vacuum is 3.00×10^8 m s^{-1}, calculate the refractive index of crown glass.

By definition, the refractive index of a vacuum is exactly 1, since $n = \frac{c}{c} = 1$. Similarly, the refractive index of air is effectively equal to 1 because the speed of light in air is practically the same as its speed in a vacuum.

The definition of refractive index allows you to determine changes in the speed of light as it moves from one medium to another.

Since $n = \frac{c}{v}$, therefore $c = nv$. This applies for any material, therefore:

$$n_1 v_1 = n_2 v_2$$

where n_1 is the refractive index of the first material

v_1 is the speed of light in the first material

n_2 is the refractive index of the second material

v_2 is the speed of light in the second material

Worked example 9.1.3

SPEED OF LIGHT CHANGES

A ray of light travels from crown glass (n = 1.52), where it has a speed of 1.97×10^8 m s^{-1} into water (n = 1.33). Calculate the speed of light in water.

Thinking	Working
Recall the formula.	$n_1v_1 = n_2v_2$
Substitute the appropriate values into the formula and solve.	$1.52 \times 1.97 \times 10^8 = 1.33 \times v_2$ $\frac{1.52 \times 1.97 \times 10^8}{1.33} = v_2$ $v_2 = 2.25 \times 10^8$ m s^{-1}

Worked example: Try yourself 9.1.3

SPEED OF LIGHT CHANGES

A ray of light travels from water (n = 1.33), where it has a speed of 2.25×10^8 m s^{-1}, into glass (n = 1.85). Calculate the speed of light in glass.

Snell's law

Refractive indexes can also be used to determine how much a light ray will refract as it moves from one medium to another. Consider the situation shown in Figure 9.1.8, where light refracts as it moves from air into water.

In 1621, the Dutch mathematician Willebrord Snell described the geometry of this situation with a formula now known as **Snell's law**:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

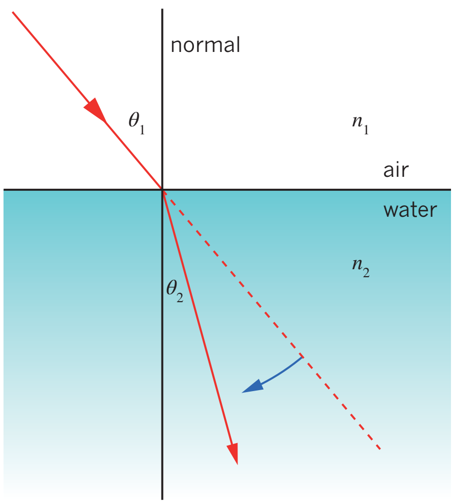


FIGURE 9.1.8 Light refracts as it moves from air into water.

Worked example 9.1.4

USING SNELL'S LAW

A ray of light in air strikes the surface of a pool of water (n = 1.33) at angle of 30° to the normal. Calculate the angle of refraction of the light in water.

Thinking	Working
Recall Snell's law.	$n_1 \sin \theta_1 = n_2 \sin \theta_2$
Recall the refractive index of air.	$n_1 = 1.00$
Substitute the appropriate values into the formula to find a value for $\sin \theta_2$.	$1.00 \times \sin 30° = 1.33 \times \sin \theta_2$ $\sin \theta_2 = \frac{1.00 \times \sin 30°}{1.33} = 0.3759$
Calculate the angle of refraction.	$\theta_2 = \sin^{-1} 0.3759 = 22.1°$

Worked example: Try yourself 9.1.4

USING SNELL'S LAW

A ray of light in air strikes a piece of flint glass (n = 1.62) at angle of incidence of 50° to the normal. Calculate the angle of refraction of the light in the glass.

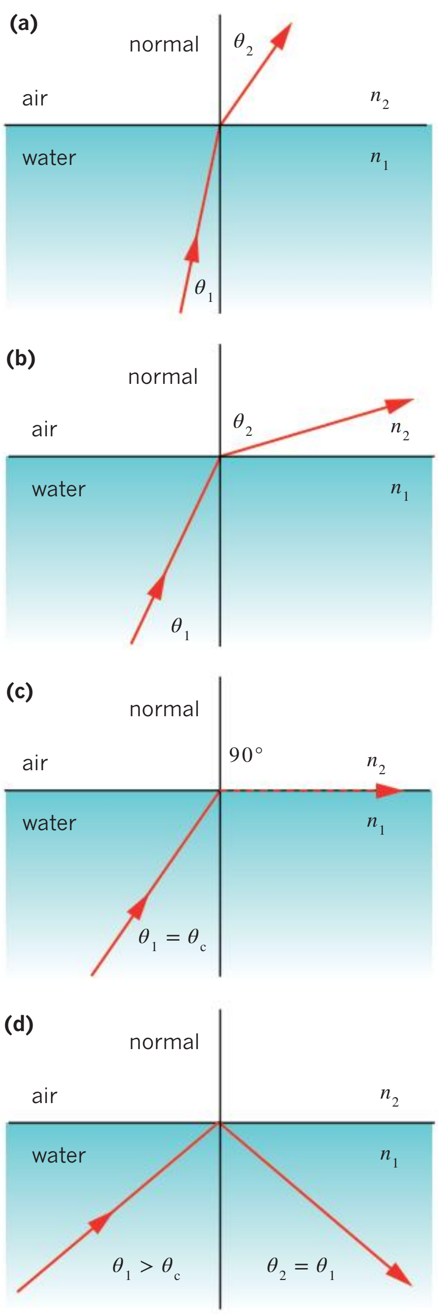


FIGURE 9.1.9 Light refracts as it moves from air into water as shown in diagrams (a) and (b). In diagram (d) the light is undergoing total internal reflection.

Total internal reflection

When light passes from a medium with low refractive index to one with higher refractive index, it is refracted *towards* the normal. Conversely, as shown in Figure 9.1.9, when light passes from a medium with a high refractive index to one with a lower refractive index, it is refracted *away from* the normal (see Figure 9.1.9(a)). In this case, as the angle of incidence increases, the angle of refraction gets closer to 90° (see Figure 9.1.9(b)). Eventually, at an angle of incidence known as the **critical angle**, the angle of refraction becomes 90° and the light is refracted along the interface between the two mediums (Figure 9.1.9(c)). If the angle of incidence is increased above this value, the light ray does not undergo refraction; instead it is reflected back into the original medium, as if it was striking a perfect mirror (see Figure 9.1.9(d)). This phenomenon is known as **total internal reflection** and is seen in action in Figure 9.1.10.

Since the angle of refraction for the critical angle is 90°, the critical angle is defined by the formula:

$n_1 \sin \theta_c = n_2 \sin 90°$

Since $\sin 90° = 1$, then $n_1 \sin \theta_c = n_2$, or $\sin \theta_c = \frac{n_2}{n_1}$.

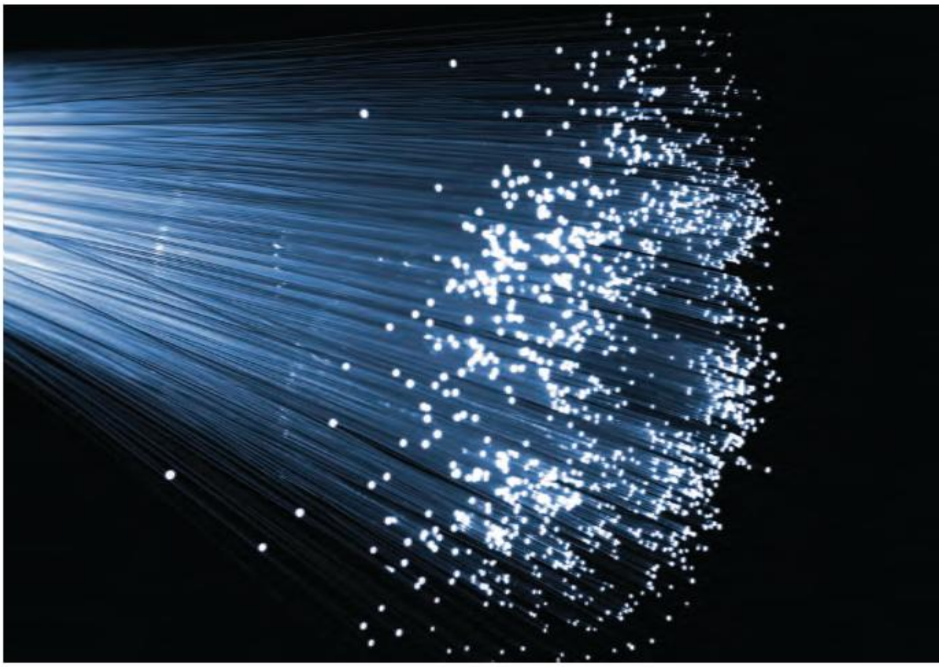


FIGURE 9.1.10 Optical fibres transmit light using total internal reflection.

Worked example 9.1.5

CALCULATING CRITICAL ANGLE

Calculate the critical angle for light passing from water into air.	
Thinking	**Working**
Recall the equation for critical angle.	$\sin \theta_c = \frac{n_2}{n_1}$
Substitute the refractive indices of water and air into the formula. (Unless otherwise stated, assume that the second medium is air with $n_2 = 1$.)	$\sin \theta_c = \frac{1.00}{1.33} = 0.7519$
Solve for θ_c.	$\theta_c = \sin^{-1} 0.7519 = 48.8°$

Worked example: Try yourself 9.1.5

CALCULATING CRITICAL ANGLE

Calculate the critical angle for light passing from diamond into air.

DISPERSION

When white light passes through a triangular glass prism (as shown in Figure 9.1.12) it undergoes **dispersion**. This is a result of refraction.

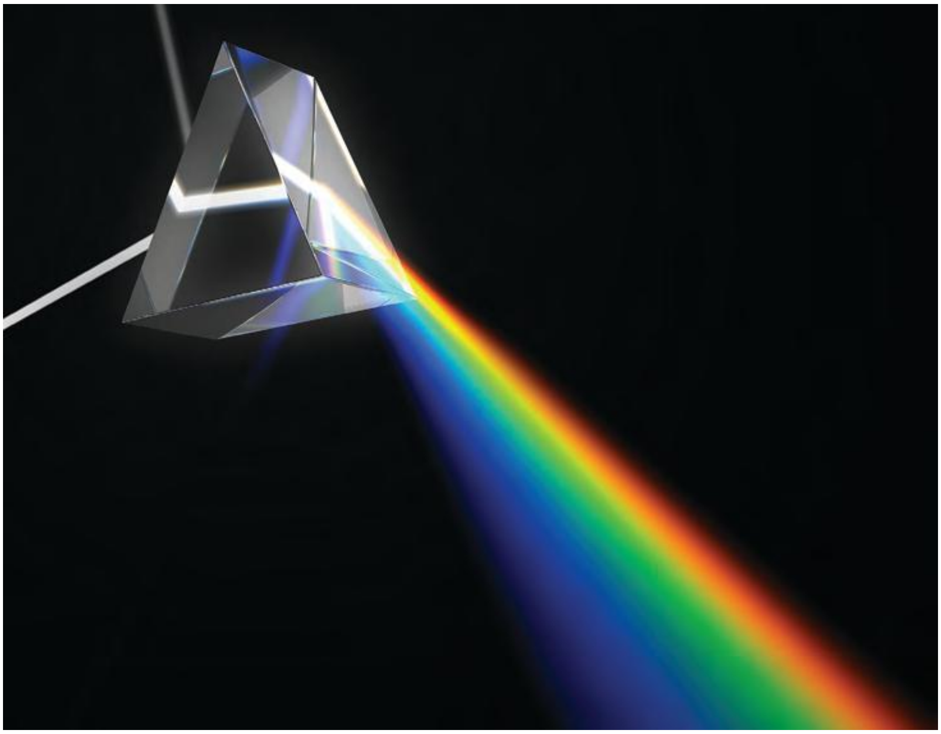


FIGURE 9.1.12 The wave model of light can explain the phenomenon of dispersion of light; the splitting of white light into its component colours.

According to the wave model of light, each different colour represents a wave of a different wavelength (see Table 9.1.3). White light is a mixture of light waves with many different wavelengths.

Colour	Wavelength (nm)
red	780–622
orange	622–597
yellow	597–577
green	577–492
blue	492–455
violet	455–390

TABLE 9.1.3 Approximate wavelength ranges for the colours in the visible spectrum. 1 nm = 10^{-9} m

When white light passes from one material to another and the light waves slow down, the wavelength shortens as the waves bunch up and the wavelengths of each colour change by different amounts. This means that each colour travels at a slightly different speed in the new medium and therefore each colour is refracted by a slightly different amount.

Longer wavelengths, such as those in red light, travel the fastest in the new material so they are refracted the least. Shorter wavelengths, such as those in violet light, are slower so they are refracted the most.

So in effect, each colour of light has a different refractive index in a material.

PHYSICSFILE

Refractive index of diamonds

Since diamond has a very high refractive index, it has a small critical angle. This means that a light ray that enters a diamond will often bounce around inside the diamond many times before leaving the diamond. A jeweller can cut a diamond to take advantage of this property; this causes the diamond to 'sparkle' (see Figure 9.1.11) as it appears to reflect more light than is falling on it.

FIGURE 9.1.11 The refractive properties of diamonds mean they appear to sparkle.

PHYSICSFILE

Where does colour come from?

In the 17th century, many people believed that white light was ‘stained’ by its interaction with earthly materials. Newton very neatly disproved this with a simple experiment using two prisms (as seen in Figure 9.1.13)—one to split light into its component colours and the other to turn it back into white light. This showed that the various colours were intrinsic components of white light since, if colour was a result of ‘staining’, the second prism should have added more colour rather than less.

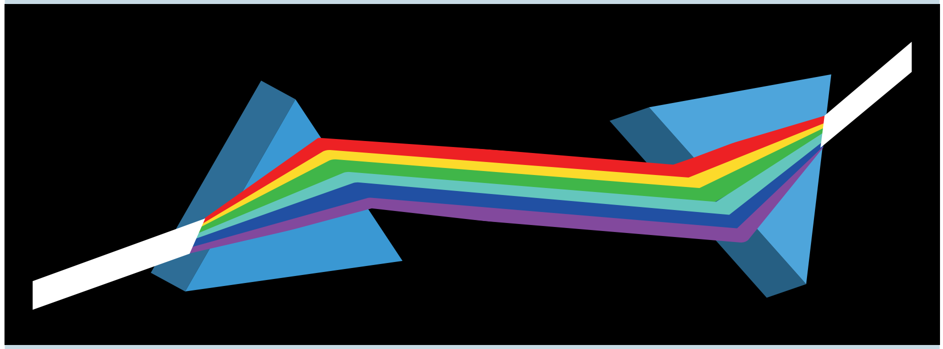


FIGURE 9.1.13 Newton’s double prism experiment showed that white light is made up of its component colours.

Newton was the first to identify the colours of the spectrum—red, orange, yellow, green, blue, indigo and violet. He chose seven colours by inventing the colour ‘indigo’ because seven was considered a sacred number.

Colour dispersion in lenses

Since each colour of light effectively has a different refractive index in glass, light passing through a glass lens always undergoes some dispersion. This means that coloured images formed by optical instruments such as microscopes and telescopes can suffer from a type of distortion known as *chromatic aberration* (see Figure 9.1.14).

FIGURE 9.1.14 Chromatic aberration causes the coloured fringes that can be seen in the circled regions in this image.

Scientists have developed a number of techniques to deal with this problem, including:

- using lenses with very long focal lengths
- using 'achromatic' lenses. These are compound lenses which are made of different types of glass with different refractive properties
- taking separate images using coloured filters and then combining these images to form a single multi-coloured image.

DIFFRACTION

When a plane (straight) wave passes through a narrow opening, it bends. Waves will also bend as they travel around obstacles (see Figure 9.1.15). This kind of 'bending' phenomenon is known as **diffraction**.

Diffraction is significant when the size of the opening or obstacle is similar to or smaller than the wavelength of the wave. This means that the diffraction of light is difficult to observe because the wavelength of light is very small.

Light waves range in wavelength from around 700 nm for red light to about 400 nm for violet light. 1 nm is equal to 10^{-9} m or a one millionth of a millimetre. This means that light waves are all less than one thousandth of a millimetre in length.

Since there are not many situations where light encounters naturally occurring structures of this size, diffraction of light waves is not easily observed.

Usually diffraction occurs with artificially constructed materials like CDs or commercially produced diffraction gratings (see Figure 9.1.16).

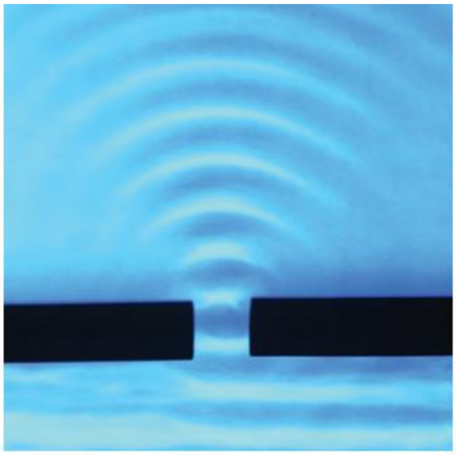


FIGURE 9.1.15 Water waves will bend around an obstacle. Sound waves diffract as well, allowing you to hear around corners.

FIGURE 9.1.16 The way information is printed on a CD or DVD means that it creates structures small enough to cause light to diffract.

Diffraction and slit width

In the diffraction of waves, if the wavelength is much smaller than the gap or obstacle, the degree of diffraction is less. For example, Figure 9.1.17 shows the diffraction of water waves in a ripple tank. In Figure 9.1.17(a), the gap is similar in size to the wavelength, so there is significant diffraction and the waves emerge as circular waves. In Figure 9.1.17(b), the gap is much bigger than the wavelength, so diffraction only occurs at the edges.

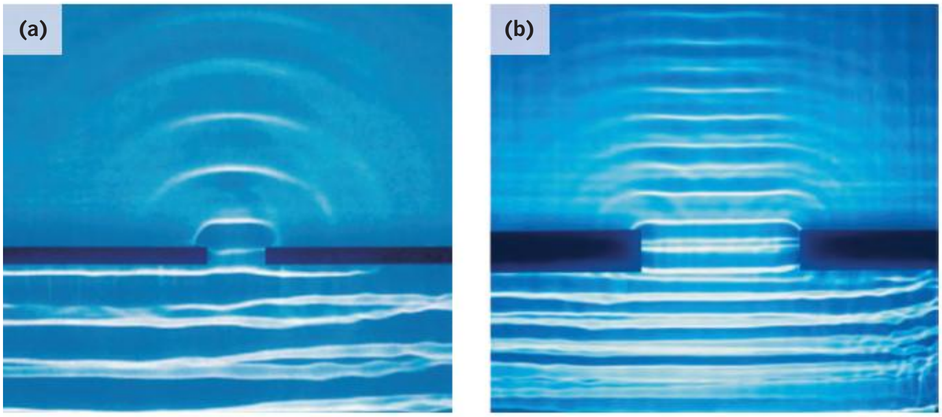


FIGURE 9.1.17 The diffraction of water waves in a ripple tank. (a) Significant diffraction occurs when the wavelength approximates the slit width, i.e. $\lambda \approx w$. (b) As the gap increases, diffraction becomes less obvious, since $\lambda << w$, but is still present.

Wavelengths comparable to or larger than the diameter of the obstacle or gap will produce significant diffraction. This can be expressed as the ratio $\frac{\lambda}{w} \geq 1$, where λ is the wavelength of the wave and w is the width of the gap.

Diffraction and imaging

Diffraction can be a problem for scientists using microscopes and telescopes because it can result in blurred images. For example, a significant problem is that the light from two tiny objects or two distant objects very close together can be diffracted so much that the two objects appear as one blurred object. When this happens, we say that the objects are unresolved. Essentially, the ratio $\frac{\lambda}{w}$ dictates how small an object can be clearly imaged by a particular instrument.

This means that, as a general rule, optical microscopes cannot create images of objects that are smaller than the wavelength of the light they use; otherwise, diffraction effects are too significant.

Diffraction also places a theoretical limit on the resolution of optical telescopes. However, atmospheric distortion usually has a much larger effect on telescope images than diffraction. The Hubble Space Telescope, which sits above Earth's atmosphere, is not affected by atmospheric distortion. It can resolve images right down to its diffraction limit, i.e. where the separation of the stars is approximately equal to the wavelength of the light.

Diffraction gratings

As you have already seen, light diffracts as it passes through a very small gap. As the light passes through the gap, some of the wavelets making up the wavefront will diffract at the barriers that form the edges of the gap and some will pass through the centre of the gap. As a result of this the light waves that emerge from the gap will interact. In some places the interactions will be constructive and in others the interactions will be destructive. When these light waves are made to shine on a screen, the areas of **constructive interference** will appear as bright bands and areas of **destructive interference** will appear as dark bands. The pattern of dark and light bands that is seen when light passes through a single small gap is called a **diffraction pattern**.

As stated earlier, the extent of diffraction of light waves is proportional to the ratio $\frac{\lambda}{w}$. This ratio also describes the spacing of dark and light bands in a diffraction pattern, and therefore the width of the overall diffraction pattern. According to this relationship, if the wavelength is held constant and the gap made smaller, greater diffraction is seen. If different wavelengths enter the same gap, those with a smaller wavelength will undergo less diffraction than those with a longer wavelength. This is shown in the Figure 9.1.18. Note that Figure 9.1.18 shows intensity. High intensity is where bright bands will appear on a screen; zero intensity corresponds to dark bands.

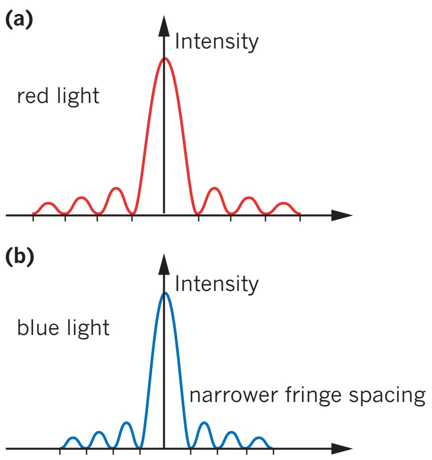


FIGURE 9.1.18 Red light (a) is diffracted to a greater extent than blue light (b). Red's longer wavelength results in more-widely spaced fringes and a wider overall pattern.

Although some diffraction patterns can be observed using natural materials, in practice, much clearer diffraction patterns can be generated by passing light through a *diffraction grating*. A diffraction grating is a piece of material that contains a large number of very closely spaced parallel gaps or slits.

A diffraction grating can be thought of as a series of parallel slits all placed side by side. The diffraction pattern from one slit is superimposed on the pattern from the adjacent slit, producing a strong, clear image on the screen.

Diffraction experiments usually use only **monochromatic** light (i.e. light of only one colour). When white light, which contains a number of different colours, shines through a diffraction grating, each different colour is diffracted by a different amount and forms its own set of coloured fringes. This results in the light being dispersed into its component colours, as seen in Figure 9.1.19.

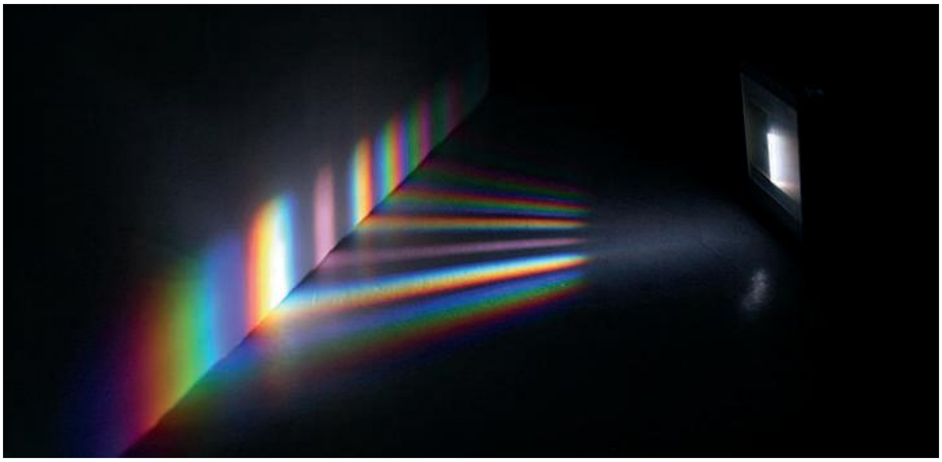


FIGURE 9.1.19 A diffraction grating disperses white light into a series of coloured spectra.

POLARISATION

One of the most convincing pieces of evidence for the wave nature of light is the phenomenon of **polarisation**. Polarisation occurs when a transverse wave (which you will recall from the previous chapter 'Properties of mechanical waves') is only allowed to vibrate in one direction. For example, the light wave in Figure 9.1.20 is vertically polarised—the wave oscillations occur in the vertical plane only. This also means this wave is unaffected by a polarising filter that is oriented in the vertical plane.

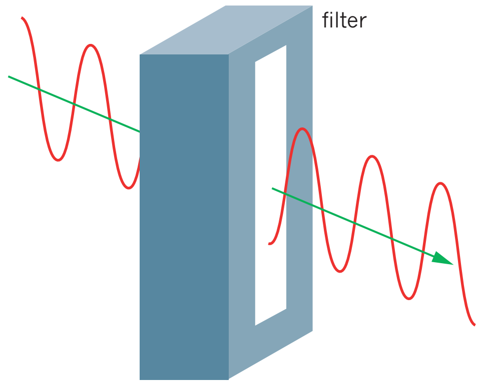


FIGURE 9.1.20 A vertically polarised wave can pass through a vertically oriented polarising filter.

The wave in the Figure 9.1.21 is horizontally polarised. It is completely blocked by the vertical polarising filter.

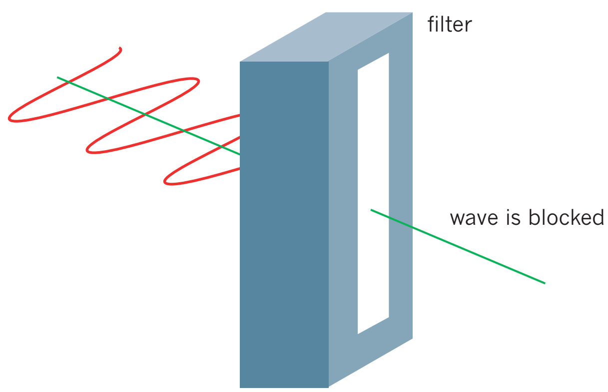


FIGURE 9.1.21 A horizontally polarised wave cannot pass through a vertically oriented polarising filter.

In Figure 9.1.22, the incoming wave is polarised at 45° to the horizontal and vertical planes. The horizontal component of this wave is blocked by the vertical filter, so the ongoing wave is vertically polarised and has a smaller amplitude than the original wave.

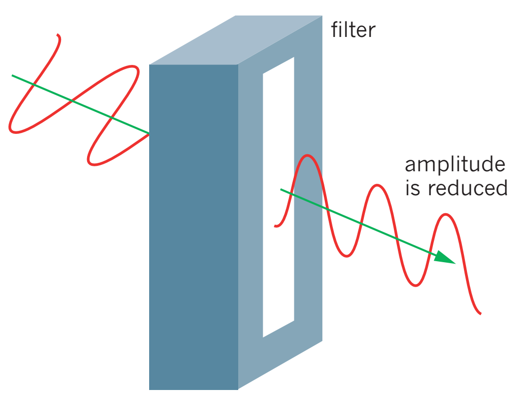


FIGURE 9.1.22 A diagonally polarised wave has its horizontal component suppressed by the vertically oriented polarising filter. A vertically polarised wave of reduced amplitude passes through it.

Light produced by sources such as a light globe or the Sun is unpolarised, which means that it can be thought of as a collection of waves, each with a different plane of polarisation, as shown in Figure 9.1.23.

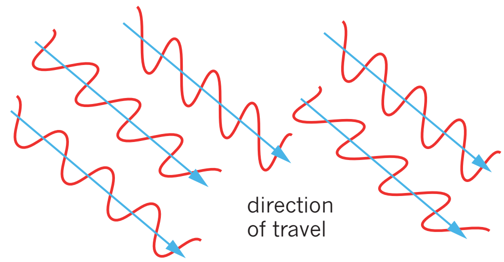


FIGURE 9.1.23 Unpolarised light consists of a collection of waves that are each polarised in a different direction.

Certain materials can act as polarising filters for light. These only transmit the waves or components of waves that are polarised in a particular direction and absorb the rest. Polarising sunglasses work by absorbing the light polarised in a particular direction, thus reducing glare. Photographers also use polarised filters to reduce the glare in photographs or achieve specific effects (see Figure 9.1.24).

FIGURE 9.1.24 These are photos taken of the same tree, one without a polarising filter (left) and one with a polarising filter (right).

PHYSICSFILE

Polarised sunglasses

Light that is reflected from the surface of water or snow is partially polarised (see Figure 9.1.25). The polarising plane of polarised sunglasses is selected to absorb this reflected light. This makes polarised sunglasses particularly effective for people involved in outdoor activities such as boating, fishing or skiing.

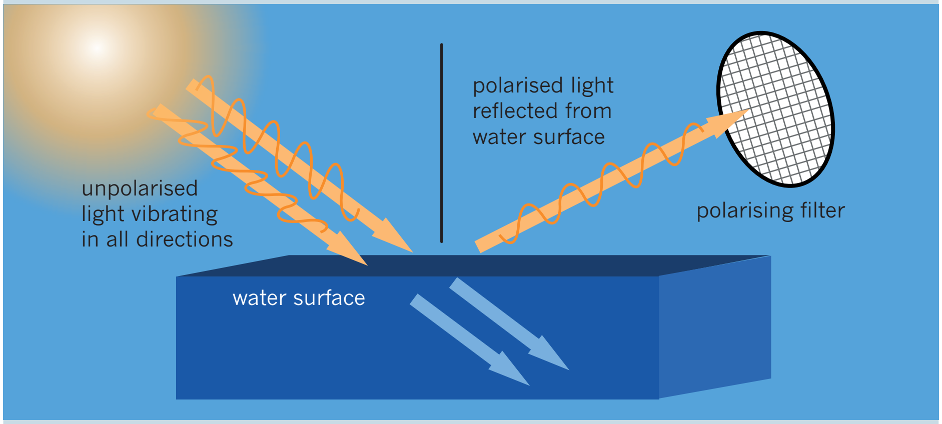


FIGURE 9.1.25 The polarising filter in a pair of sunglasses is designed to block the polarised light reflected from the surface of the water and transmit the unpolarised light from under the water.

9.1 Review

SUMMARY

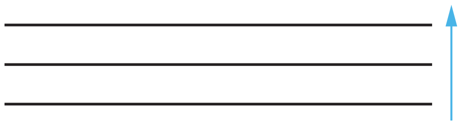
- A wave model explains a wide range of light-related phenomena, including refraction, dispersion, diffraction and interference.
- Refraction is the change in the direction of light that occurs when light moves from one medium to another.
- Refraction is caused by changes in the speed of light waves.
- The refractive index, n, of a material is given by the formula $n = \frac{c}{v}$ where c is the speed of light in a vacuum and v is the speed of light in the material.
- When light moves from one material to another, the changes in speed can be calculated using:
  $n_1 v_1 = n_2 v_2$
- The amount of refraction of a ray of light can be calculated using Snell's law: $n_1 \sin \theta_1 = n_2 \sin \theta_2$
- Total internal reflection occurs when the angle of refraction exceeds a right angle.
- The critical angle of a material can be calculated using $n_1 \sin \theta_c = n_2 \sin 90°$ or $\sin \theta_c = \frac{n_2}{n_1}$
- Different colours of light have different wavelengths.
- Dispersion occurs because different colours of light travel at different speeds.
- When a plane (straight) wave passes through a narrow opening or meets a sharp object, it experiences diffraction.
- Significant diffraction occurs when the wavelength of the wave is similar to, or larger than, the size of the diffracting object.
- A transverse wave model is required to explain polarisation.

KEY QUESTIONS

1 Name the model of light each of the following scientists supported.
   a Hooke
   b Huygens
   c Newton

2 In the 18th century, why did most scientists support Newton's particle model?
   A Newton had better evidence to support his theory.
   B The speed of light in glass had been shown to be faster than in air.
   C Newton had a better reputation as a scientist than either Hooke or Huygens.
   D Newton was English and Hooke and Huygens were from other parts of Europe.

3 Draw the wavefront of the plane wave after one period.



4 Choose the correct response from those given in bold to complete the sentences about the refractive indicies of types of water.
   Although pure water has a refractive index of 1.33, the salt content of seawater means its refractive index is a little higher at 1.38. Therefore, the speed of light in seawater will be **faster than/slower than/the same as** in pure water.

5 Calculate the speed of light in seawater that has a refractive index of 1.38.

6 Light travels at of 2.25×10^8 m s^{-1} in water and 2.29×10^8 m s^{-1} in ice. If water has a refractive index of 1.33, use this information to calculate the refractive index of ice.

7 Light travels from water (n = 1.33) into glass (n = 1.60). The incident angle is 44°. Calculate the angle of refraction.

8 For which of the following situations can total internal reflection occur?

	Incident medium	Refracting medium
a	air (n = 1.00)	glass (n = 1.55)
b	glass (n = 1.55)	air (n = 1.00)
c	glass (n = 1.55)	water (n = 1.33)
d	glass (n = 1.55)	glass (n = 1.58)

9 In order to produce significant diffraction of red light (wavelength of approximately 700 nm), a diffraction experiment would need to use an opening with a width of approximately:
   A 1 mm
   B 0.1 mm
   C 0.01 mm
   D 0.001 mm

10 Explain how polarisation supports a wave model for light.

9.2 Interference: Further evidence for the wave model of light

Thomas Young's observation of interference patterns in light (see Figure 9.2.1) was a pivotal moment in the history of science. It tipped the scales in a long-running dispute between scientists about the nature of light and paved the way for a series of discoveries and inventions that would fundamentally change scientists' understanding of energy and matter.

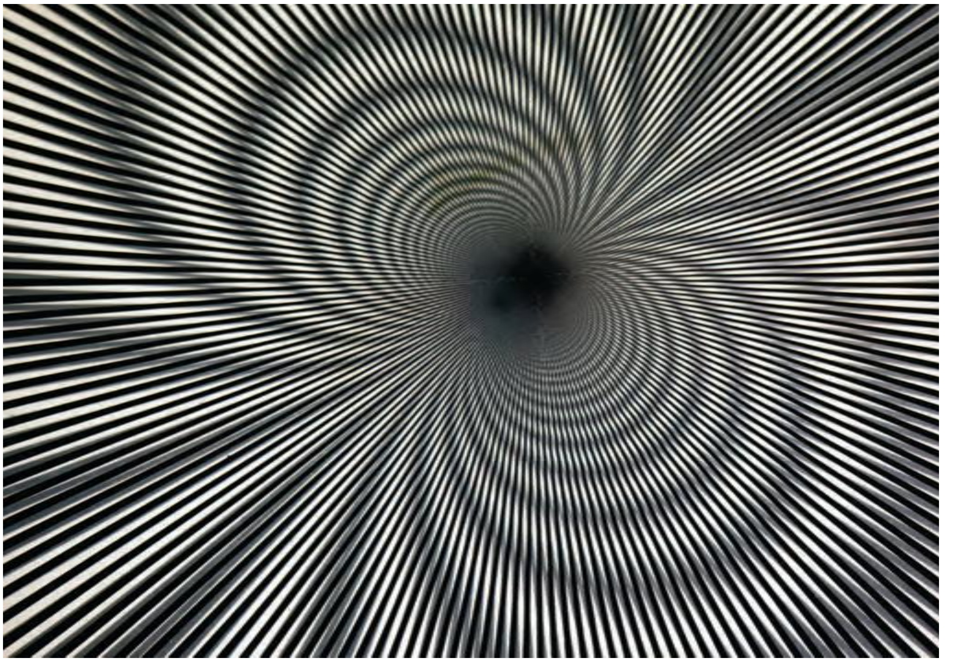


FIGURE 9.2.1 Optical interference can produce spectacular patterns.

YOUNG'S DOUBLE-SLIT EXPERIMENT

Between the 17th and 19th centuries, most scientists considered light to be a stream of particles. This idea was based on the 'corpuscular' theory proposed by Sir Isaac Newton.

In 1803, an English scientist called Thomas Young performed a now-famous experiment in which he shone monochromatic light on a screen containing two very tiny slits. On the far side of the double slits he placed another screen, on which he observed the pattern produced by the light passing through the slits (see Figure 9.2.2).

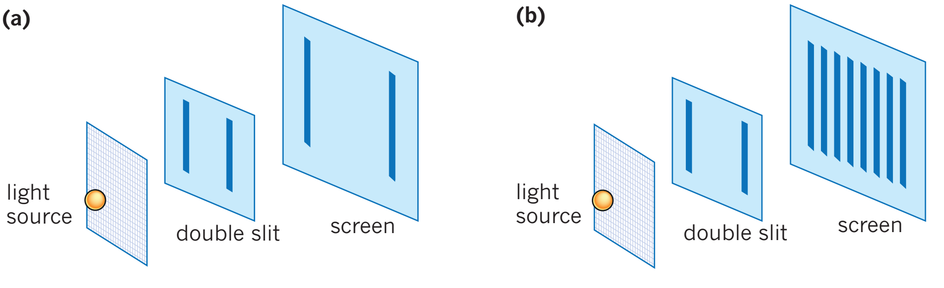


FIGURE 9.2.2 The particle theory of light predicted that Young's experiment should produce two bright bands (a). But his actual experiment (b) produced a series of bright and dark bands or 'fringes'.

According to the particle theory, light should have passed directly through the slits to produce two bright lines or bands on the screen (see Figure 9.2.2(a)). Instead, Young observed a series of bright and dark bands or 'fringes' (see Figure 9.2.2(b)).

Young was able to explain this bright and dark pattern by treating light as a wave. He assumed that the monochromatic light was like plane waves and that, as they passed through the narrow slits, these plane waves were diffracted into **coherent** (in phase) circular waves as shown in Figure 9.2.3. The circular waves would interact causing **interference**. The interference pattern produced by these two waves would result in lines of constructive (antinodal) and destructive (nodal) interference that would match the bright and dark fringes respectively.

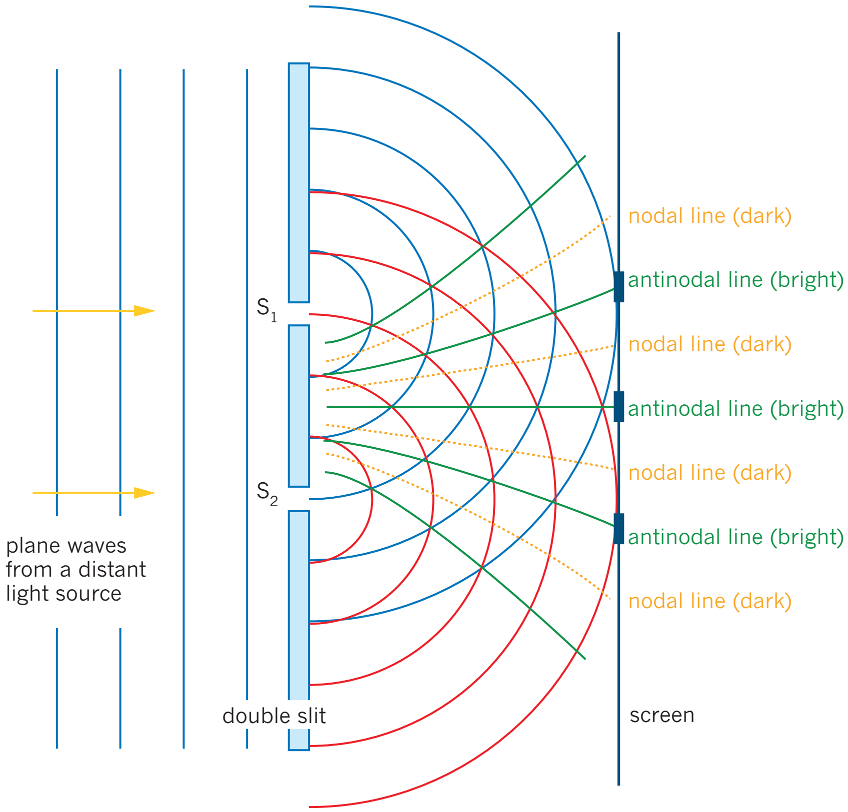


FIGURE 9.2.3 The interaction of two circular waves can produce a pattern of antinodal (constructive interference) and nodal (destructive interference) lines.

Earlier in his scientific career, Young had observed similar interference patterns in water waves (see Figure 9.2.4). This gave greater credibility to the wave model for light proposed by Christiaan Huygens and Robert Hooke many years earlier.

When Young used his data to calculate the wavelength of light, it became clear why no one had ever noticed the wave properties of light before—light waves are tiny, with typical wavelengths of less than 1 micrometre (1 μm = 0.001 mm).

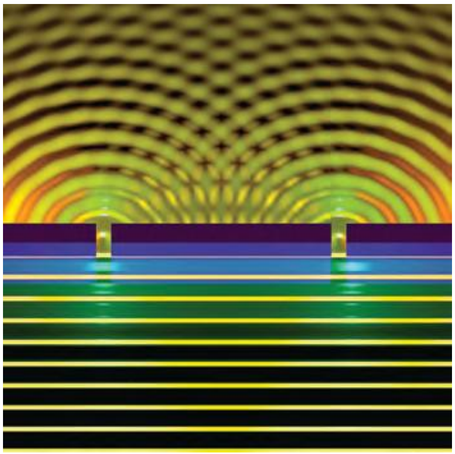


FIGURE 9.2.4 Interference patterns can be observed in water waves (lit here in yellow).

Path difference

To understand Young's experiment more fully, you have to consider how the waves produced by the two slits interact with each other when they hit the screen. At a particular point, P, on the screen, the wave train from slit 1 (S_1) will have travelled a different distance compared with the wave train from slit 2 (S_2), i.e. the distance S_1P Is different to S_2P. The difference in the distance travelled by each wave train to a point P on the screen is called the **path difference** for the waves (pd).

The path difference to point P from wave source S_1 and from wave source S_2 is given by:

$$pd = |S_1P - S_2P|$$

Path difference can be measured in metres, but it is far more useful to measure it in wavelengths in order to determine the light intensity on the screen.

As shown in Figure 9.2.5, at a point, M, at the centre of the screen, equidistant from each slit, each wave train will have travelled through the same distance and so there is no path difference (i.e. $S_1M = S_2M$). The light waves arrive in phase with each other. These light waves reinforce to produce an antinode. A fringe of bright light is seen, known as the 'central maximum'. This phenomenon is called constructive interference.

Constructive interference will occur whenever the path difference between the two wave trains is zero or differs by a whole number of wavelengths, i.e. pd = 0λ, 2λ, 3λ... For example, in Figure 9.2.5, the path difference $S_1R - S_2R$ is equal to λ.

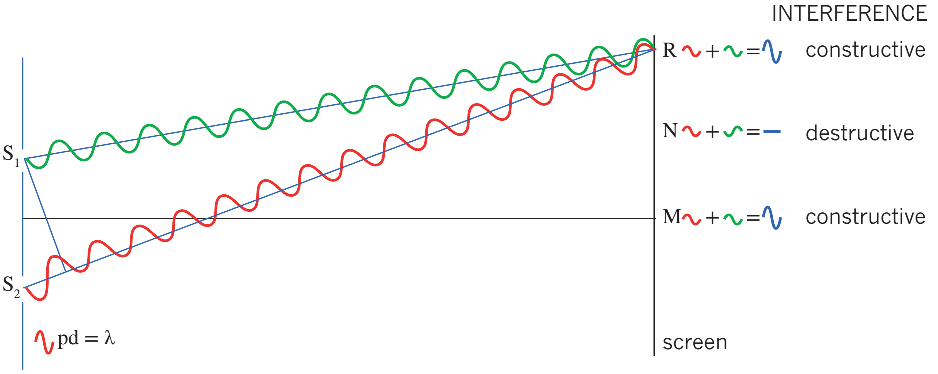


FIGURE 9.2.5 Waves meeting from each slit at R, where the path difference is λ. A bright fringe will be seen as the wave trains arrive at this point in phase again.

There will be points on the screen at which the path difference is $\frac{\lambda}{2}$; for example point N in Figure 9.2.5. The two wave trains that meet at this point are completely out of phase and cancel each other to produce a nodal point. Destructive interference occurs at this point, and no light is seen. This creates the dark lines or fringes that appear in between the bright antinodal fringes. Destructive interference occurs when the path difference between the waves is $\frac{\lambda}{2}$, $\frac{3\lambda}{2}$, $\frac{5\lambda}{2}$, $\frac{7\lambda}{2}$, etc.

In summary:

- constructive interference of coherent waves occurs when the path difference pd = $n\lambda$, where n = 0, 1, 2, 3...
- destructive interference of coherent waves occurs when the path difference equals an odd number of half wavelengths; that is, pd = $(n - \frac{1}{2})\lambda$, where n = 1, 2, 3...

The sequence of constructive and destructive interference effects produces an interference pattern of regularly spaced vertical bands or fringes on the screen that can be represented graphically as shown in Figure 9.2.6.

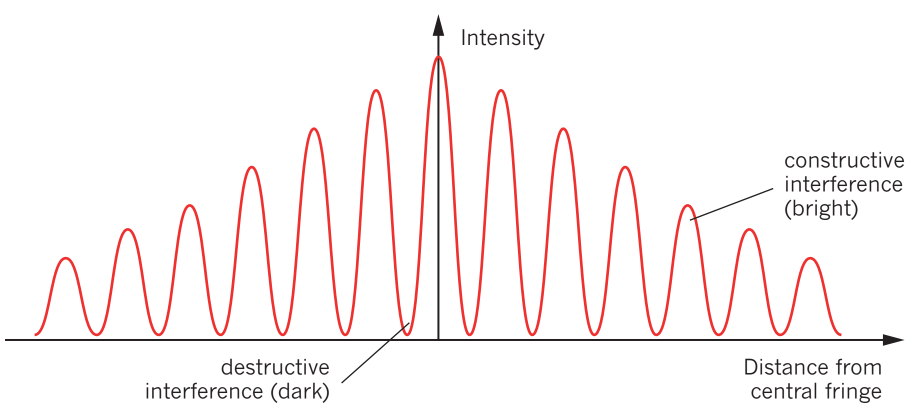


FIGURE 9.2.6 The double-slit interference pattern can be considered in terms of an intensity distribution graph. The horizontal axis represents a line drawn across the screen. The centre of the distribution pattern corresponds to the centre of the brightest central fringe, the central maximum.

Calculating fringe separation for Young's experiment

In Young's experiment, the distance between adjacent bright bands on the screen is known as the fringe spacing (Δx). This distance depends on the wavelength of light (λ), the separation between the two slits (d) and the distance to the screen (L), as shown in Figure 9.2.7.

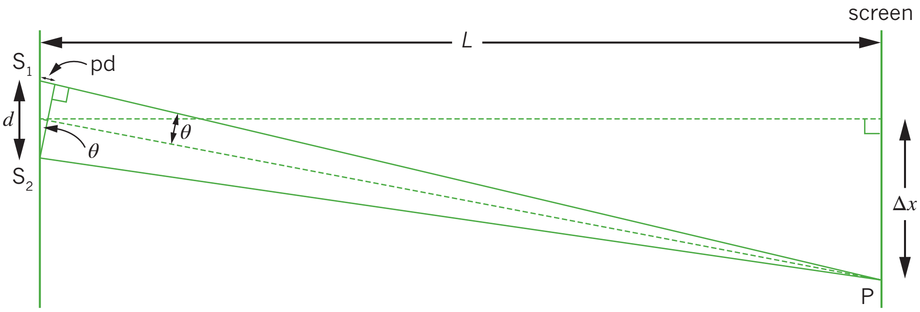


FIGURE 9.2.7 The geometry of two-point source interference.

- If the viewing screen is moved further from the two slits, the fringes will appear further apart from each other, i.e. $\Delta x \propto L$.
- Conversely, reducing the separation of the slits increases the spacing of the fringes, i.e. $\Delta x \propto \frac{1}{d}$.
- Using light of a longer wavelength will result in increased fringe spacing, i.e. $\Delta x \propto \lambda$ (see Figure 9.2.8).

Fringe separation parameters

These relationships can be combined to develop an overall equation for the fringe separation:

$$\Delta x = \frac{\lambda L}{d}$$

where Δx is the fringe separation

λ is the wavelength of the light waves

L is the distance from the slits to the screen

d is the slit separation

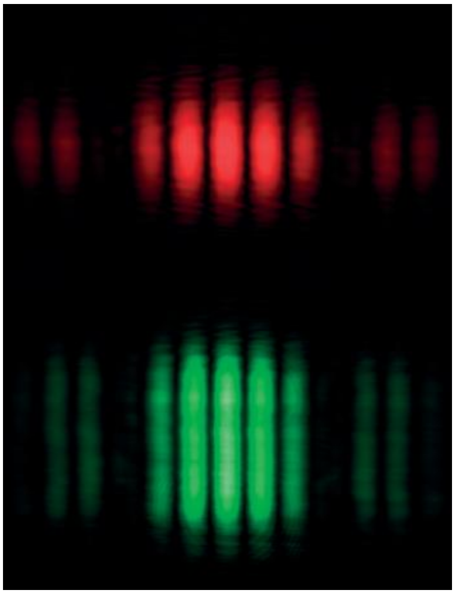


FIGURE 9.2.8 If the separation of the slits and the distance to the screen are kept the same, then the fringes produced by longer wavelength red light are further apart than for shorter wavelength green light.

Worked example 9.2.1

CALCULATING WAVELENGTH FROM FRINGE SEPARATION

Light of an unknown wavelength emitted by a laser is directed through a pair of thin slits separated by 50 μm. The slits are 2.0 m from a screen on which bright fringes are 2.5 cm apart. Calculate the wavelength of the laser light in nm.	
Thinking	**Working**
Recall the equation for fringe separation.	$\Delta x = \frac{\lambda L}{d}$
Transpose the equation to make λ the subject.	$\lambda = \frac{\Delta x d}{L}$
Substitute values into the equation and solve. (Note: 1 μm = 1×10^{-6} m)	$\lambda = \frac{0.025 \times 50 \times 10^{-6}}{2.0} = 6.25 \times 10^{-7}$ m
Express your answer using convenient units—in this case nm, where 1 nm = 1×10^{-9} m	The wavelength of the laser light is 625 nm.

Worked example: Try yourself 9.2.1

CALCULATING WAVELENGTH FROM FRINGE SEPARATION

Green laser is directed through a pair of thin slits that are 25 μm apart. The slits are 1.5 m from a screen on which bright fringes are 3.3 cm apart. Use this information to calculate the wavelength of green light in nm.

RESISTANCE TO THE WAVE MODEL

Young's wave explanation for his experiment was not immediately accepted by the scientific community. Many scientists were reluctant to abandon the corpuscular theory that had been accepted for over a century.

In 1818, the French scientist Augustin-Jean Fresnel was able to provide a mathematical explanation for Young's double-slit experiment based on Huygens' principle.

Another French scientist, Simeon Poisson, who was a passionate supporter of Newton's particle theory, argued that if the same mathematics was applied to the light shining around a round disk, then there should be a bright spot in the middle of the shadow created by the disk (see Figure 9.2.9). Since nobody had ever observed a bright spot in the middle of a shadow, Poisson believed this proved that the wave model was incorrect.

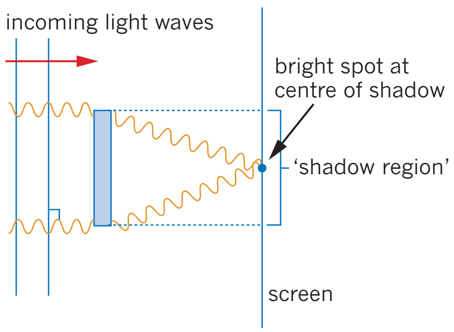


FIGURE 9.2.9 Waves of light incident on a solid disk diffract to give a point of light in the centre of the shadow zone. This is convincing evidence for the wave nature of light.

However, one of Poisson's colleagues decided to test these ideas by performing an experiment with a very small bright light source and a round disk, and observed the bright spot predicted by Poisson's calculations (see Figure 9.2.10). As a consequence, for the remainder of the 19th century, the wave theory became the almost universally accepted model for light.

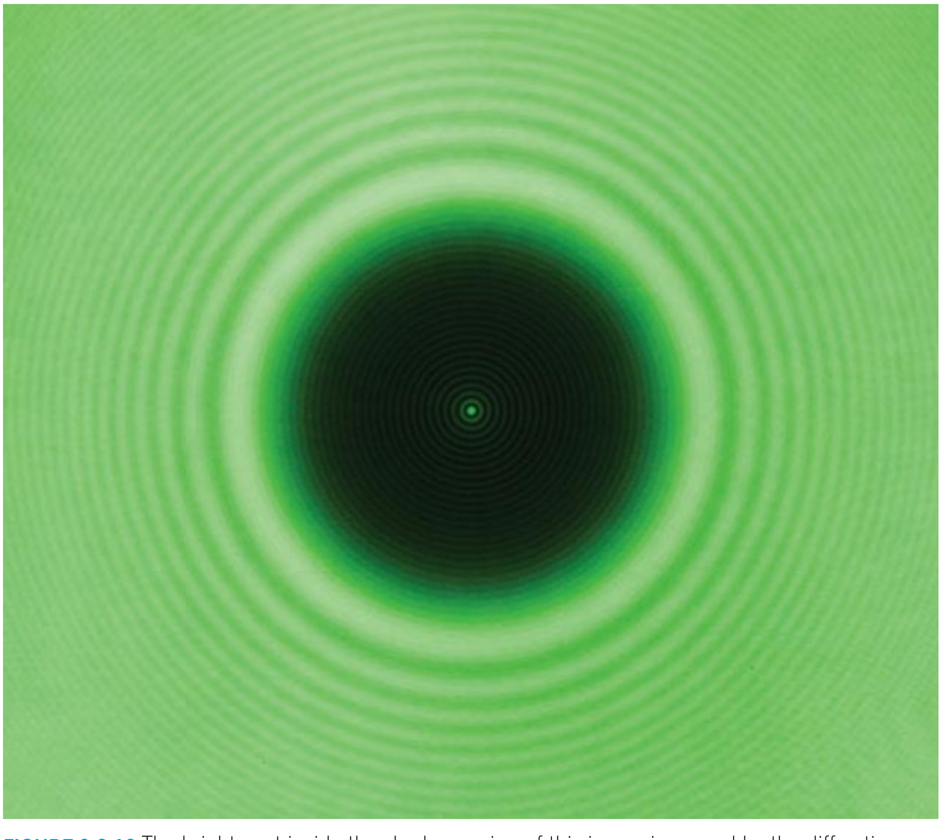


FIGURE 9.2.10 The bright spot inside the shadow region of this image is caused by the diffraction and interference of light waves. The image also shows diffraction and interference patterns surrounding the shadow.

This now famous diffraction pattern has come to be known as the 'Poisson bright spot', which means it is named after the person who predicted that it would not exist!

9.2 Review

SUMMARY

- Young's double-slit interference experiment provided evidence to support the wave model of light.
- Path difference (pd) is the difference in the distance travelled by each wave train from a pair of slits to the same point on the screen.
- Constructive interference of coherent waves occurs when the path difference pd = $n\lambda$, where n = 0, 1, 2, 3...
- Destructive interference of coherent waves occurs when the path difference equals an odd number of half wavelengths or pd = $(n - \frac{1}{2})\lambda$ where n = 1, 2, 3...
- The distance between the interference fringes produced in Young's experiment is given by: $\Delta x = \frac{\lambda L}{d}$

KEY QUESTIONS

1 According to the particle model of light, Young's double-slit experiment should have produced two bright lines on the screen. Instead, what was observed on the screen?

A It was completely dark.
B It was completely light.
C It contained three bright lines.
D It contained a pattern of alternating bright and dark lines.

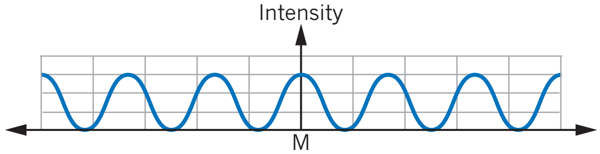
2 Two students are trying to replicate Young's double-slit experiment. One uses torch light and the other uses light from a laser. The student using the laser light is more likely to obtain the expected interference pattern because of which of the following statements (note that more than one correct answer is possible):

A torch light is monochromatic
B torch light is coherent
C laser light is monochromatic
D laser light is coherent

3 If Thomas Young's double-slit experiment was modelled using circular water waves in a ripple tank, which of the following events would correspond to nodal lines? (Note that more than one correct answer is possible.)

A crests meet troughs
B troughs meet troughs
C crests meet crests
D troughs meet crests

4 The following diagram shows the resulting (simplified) intensity pattern after light from two slits reaches the screen in a Young's interference experiment. Copy the diagram into your workbook and circle the points at which the path difference is equal to 1λ.



5 Explain why Young's double-slit experiment led to a significant change in scientists' understanding of the nature of light.

6 A version of Young's double-slit experiment is set up by directing the light from a red laser through a pair of thin slits. An interference pattern appears on the screen behind the slits. The following changes are made to the apparatus. Identify whether the distance between the interference fringes seen on the screen would increase, decrease or stay the same.

a the screen is moved further away from the slits
b a green (i.e. shorter wavelength) laser is used
c the slits are moved closer together

7 A 580 nm yellow light is directed through a pair of thin slits to produce an interference pattern on a screen. Determine the path difference of the fifth dark fringe.

8 Identify the type of interference (constructive or destructive) that corresponds to the following path differences:

a $\frac{\lambda}{2}$
b λ
c $\frac{3\lambda}{2}$

9 A 700 nm red light is directed through a pair of thin slits to produce an interference pattern on a screen. Determine the path difference of the second bright fringe.

10 A blue laser is directed through a pair of thin slits that are 40 μm apart. The slits are 3.25 m from a screen on which bright fringes are 3.7 cm apart. Use this information to calculate the wavelength of green light in nm.

9.3 Electromagnetic waves

The establishment of the wave model for light raised an important question. Scientists now wanted to know what type of waves light waves were.

Experiments on polarisation provided the important information that light must be a type of transverse wave, since polarisation does not occur for longitudinal waves. However, light is obviously different from other types of mechanical waves because it can pass through the vacuum of space between the Earth and Sun (see Figure 9.3.1).

FIGURE 9.3.1 Light cannot be a simple mechanical wave because it can travel through empty space.

ELECTROMAGNETIC WAVES

In the middle of the 19th century, the Scottish physicist James Clerk Maxwell gained a key insight into the nature of light waves. In his mathematical study of electric and magnetic effects, he realised that some of the constants in his equations closely matched the current estimates of the speed of light. Maxwell went on to develop a comprehensive theory of electromagnetism in which light is a form of **electromagnetic radiation** (EMR).

The electromagnetic nature of light

As shown in earlier chapters, electric current can be used to produce a magnetic field, and a changing magnetic field can be used to generate an electromotive force (EMF) or voltage. Maxwell put these two ideas together. He proposed that if a changing electric field is produced, for example by a charged particle moving backwards and forwards, then this changing electric field will produce a changing magnetic field at right angles to it, as shown in Figure 9.3.2.

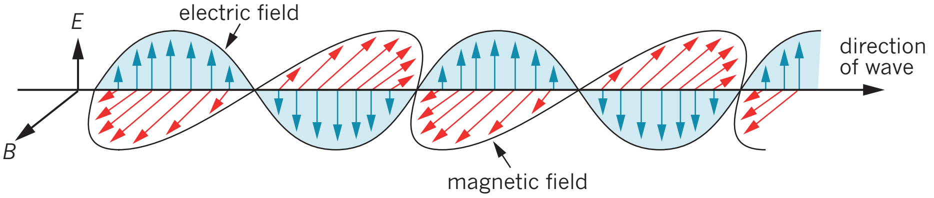


FIGURE 9.3.2 The electric and magnetic fields in electromagnetic radiation are perpendicular to each other and are both perpendicular to the direction of propagation of the radiation.

The changing magnetic field would, in turn, produce a changing electric field and the cycle would be repeated. In effect, this would produce two mutually propagating fields and the electromagnetic radiation would be self-propagating, i.e. it could extend outwards into space. Both the electric and magnetic fields would oscillate at the same frequency: the frequency of the light wave.

Maxwell's theoretical calculations provided a value for the speed at which electromagnetic radiation should propagate through empty space. This matched the experimental value for the speed of light measured by the French physicist Hippolyte Fizeau in 1849. The accepted value for the speed of light today is 299 792 458 m s^{-1}. This is such an important constant that it is designated its own symbol, c. In calculations, the speed of light is usually approximated as $c = 3.00 \times 10^8$ m s^{-1}.

For light and other forms of EMR, the familiar wave equation $v = f\lambda$ is usually written as:

$$c = f\lambda$$

where f is the frequency of the wave (Hz)

λ is the wavelength of the wave (m)

Maxwell's work represents a pivotal moment in the history of physics. Not only did he provide an explanation of the nature of light, he also brought together a number of formerly distinct areas of study—optics (the study of light), electricity and magnetism. As shown in the next section, Maxwell's work also encompasses other areas of physics.

Worked example 9.3.1

USING THE WAVE EQUATION FOR LIGHT

Calculate the frequency of violet light with a wavelength of 400 nm (i.e. 400×10^{-9} m).	
Thinking	**Working**
Recall the wave equation for light.	$c = f\lambda$
Transpose the equation to make frequency the subject.	$f = \frac{c}{\lambda}$
Substitute in values to determine the frequency of this wavelength of light.	$f = \frac{3.0 \times 10^8}{400 \times 10^{-9}}$ $= 7.5 \times 10^{14}$ Hz

Worked example: Try yourself 9.3.1

USING THE WAVE EQUATION FOR LIGHT

A particular colour of red light has a wavelength of 600 nm. Calculate the frequency of this colour.

Searching for the aether

One of the characteristics of mechanical waves is that they require a physical medium through which to propagate. For example, sound waves usually propagate through air and water waves propagate through water. For many years, scientists searched for the physical medium through which electromagnetic waves propagate. They even went so far as to give this medium a name: the 'luminiferous ether' or 'aether'.

However, all attempts to measure the presence or properties of the aether were unsuccessful. Eventually, scientists were forced to conclude that electromagnetic waves are able to propagate through a vacuum.

The electromagnetic spectrum

The wavelengths of all the different colours of visible light fall between 390 nm and 780 nm. Naturally, physicists were bound to inquire about other wavelengths of electromagnetic radiation. It is now understood that the visible spectrum is just one small part of a much broader set of possible wavelengths known as the **electromagnetic spectrum** (see Figure 9.3.3).

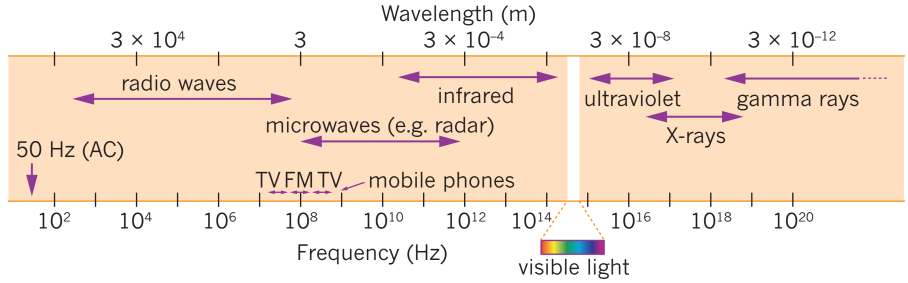


FIGURE 9.3.3 The electromagnetic spectrum.

Changing the frequency and wavelength of the waves changes the properties of the electromagnetic radiation, and so the electromagnetic spectrum is divided into 'bands' according to how the particular types of EMR are used. The shorter the wavelength of the EM wave, the greater its penetrating power. This means that waves with extremely short wavelength, such as X-rays, can pass through some materials (e.g. skin), revealing the structures inside (e.g. bone).

On the other hand, long wavelength waves such as AM radio waves have such low penetrating power that they cannot even escape Earth's atmosphere, and can be used to 'bounce' radio signals around to the other side of the world. Table 9.3.1 compares the characteristics of different waves in the electromagnetic spectrum.

Type of wave	Typical wavelength (m)	Typical frequency (Hz)	Comparable object
AM radio wave	100	3×10^{6}	sports oval
FM radio or TV wave	3	1×10^{8}	small car
microwaves	0.03	1×10^{10}	50c coin
infrared	10^{-5}	3×10^{13}	white blood cell
visible light	10^{-7}	3×10^{15}	small cell
ultraviolet	10^{-8}	3×10^{16}	large molecule
X-ray	10^{-10}	3×10^{18}	atom
gamma ray	10^{-15}	3×10^{23}	atomic nucleus

TABLE 9.3.1 Comparison of the different waves in the electromagnetic spectrum.

Radio waves

One of the most revolutionary applications of EMR is the use of radio waves to transmit information from one point to another over long distances. Radio waves are the longest type of electromagnetic radiation, with wavelengths ranging from 1 mm to hundreds of kilometres. The principle of radio transmission is relatively simple and neatly illustrates the nature of electromagnetic waves.

The radio transmitter converts the signal (e.g. radio announcer's voice, music or stream of data) into an alternating current. When this alternating current flows in the transmission antenna, the electrons in the antenna oscillate backwards and forwards. This oscillation of charges in the antenna produces a corresponding electromagnetic wave that radiates outwards in all directions from the antenna.

When the radio wave hits the antenna of a radio receiver, the electrons in the receiver's antenna start to oscillate in exactly the same way as in the transmitting antenna. The radio receiver then reverses the process of the transmitter, converting the alternating current from the reception antenna back into the original signal as seen in Figure 9.3.4.

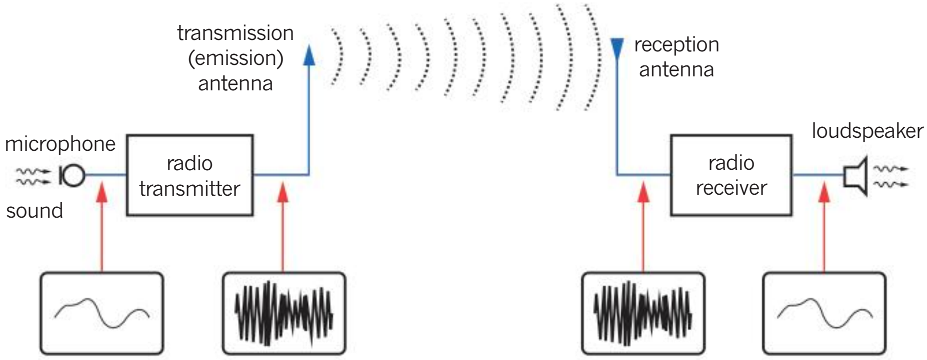


FIGURE 9.3.4 A typical radio transmission system.

PHYSICSFILE

AM and FM

A radio wave pattern is produced using a 'carrier wave' of fixed frequency. This frequency is the 'channel' that the radio 'tunes into'. Many radio stations use the carrier wave frequency as part of their name, e.g. Nova 100.3 transmits using a 100.3 MHz carrier wave.

The carrier wave is altered or 'modulated' by the signal containing the information to be transmitted. An AM radio system uses 'amplitude modulation', which means that the amplitude of the carrier wave is modulated to match the signal. In comparison, FM stands for 'frequency modulation', in which the frequency of the carrier wave is changed to represent the signal. In terms of circuitry, AM systems are much simpler than FM systems, although FM radio waves tend to transmit signals more clearly.

EXTENSION

Microwaves

Microwaves have shorter wavelengths and therefore greater penetrating power than radio waves. They can be produced by devices with short antennas and hence are useful in personal communication applications like mobile phones and wireless internet transmission. They also particularly useful in heating and cooking food (see Figure 9.3.5).

A microwave oven is 'tuned' to produce a particular frequency of electromagnetic radiation: 2.45 GHz (i.e. 2.45×10^9 Hz). This is the resonant frequency of water molecules.

All solid objects have a frequency at which they will naturally vibrate. Musical instruments such as guitars or violins make use of the resonant frequencies of strings under tension, which you might recall from Chapter 8 'Properties of mechanical waves'.

When water molecules are bombarded with radiation with a frequency of 2.45 GHz, they start to vibrate quickly. Since this increases the average kinetic energy of the water molecules, the temperature of the water in the substance increases. Effectively, the microwaves cause the water to heat up.

This heat then transfers by conduction and convection to the rest of the food. This is why food sometimes becomes soggy when heated in the microwave: the water molecules heat up faster than the food molecules around them. It also explains why recipes that do not contain much water cannot be cooked well in a microwave oven.

FIGURE 9.3.5 Microwave ovens produce electromagnetic radiation with a frequency of 2.45 GHz, which is the resonant frequency of water molecules.

Infrared

The infrared section of the electromagnetic spectrum lies between microwaves and visible light. Infrared waves are longer than the red waves of the visible spectrum, hence their name.

Infrared light waves become useful because they are emitted by objects, to varying degrees, due to their temperature (see the Physics file 'Night vision'). The warmth that you feel standing next an electric bar heater or fire is due to infrared radiation (see Figure 9.3.6). The radiant heat the Earth receives from the Sun is transmitted in the form of infrared waves; life on Earth would not be possible without this important form of electromagnetic radiation.

FIGURE 9.3.6 The coals of a fire appear red because they release red light along with infrared radiation, which you experience as heat.

PHYSICSFILE

Night vision

Night-vision goggles enhance visibility in low light conditions by greatly amplifying the visible light available and also by detecting a small part of the infrared radiation that is emitted by objects due to their temperature.

FIGURE 9.3.7 Night-vision goggles

Ultraviolet light

As their name suggests, ultraviolet (UV) waves have wavelengths that are shorter than those of violet light and therefore cannot be detected by the human eye. Their shorter wavelengths give UV rays stronger penetrating power than visible light. In fact, UV rays can actually penetrate human skin and damage the DNA of skin cells, producing harmful skin cancers.

Scientists can make use of UV light to take images. Figure 9.3.8 is a UV image of the surface of the Sun taken after a solar flare has occurred. The image has been re-coloured so that it highlights areas of different temperature. Here, areas that are coloured white are the hottest. Images like this help scientists to learn about the temperatures of very hot objects. Taking an image of the Sun using visible light would not allow this same distinction.

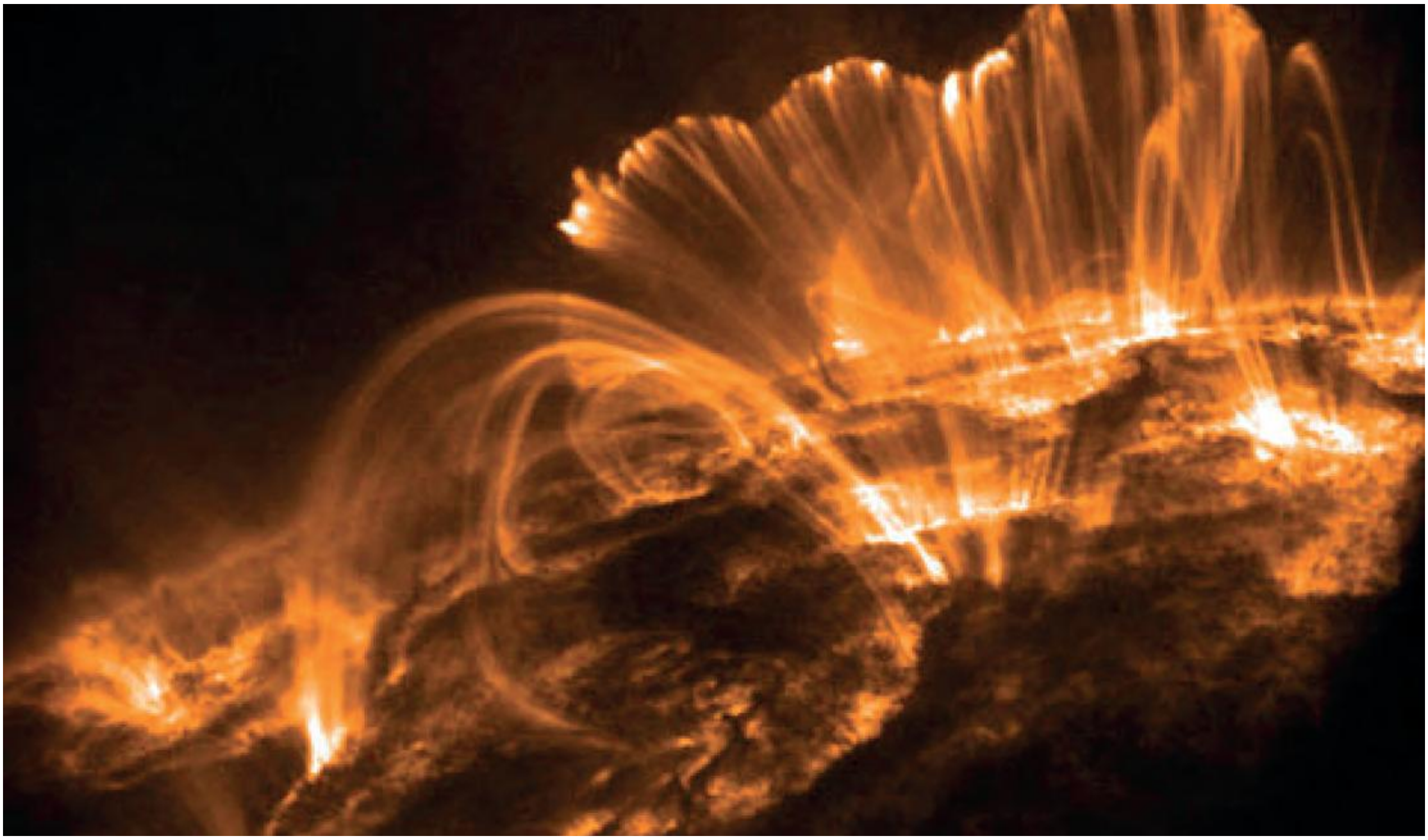


FIGURE 9.3.8 Re-coloured UV image of the surface of the Sun. The white areas reveal the hottest parts.

X-rays and gamma rays

Both X-rays and gamma rays have much shorter wavelengths than visible light. This means that these forms of electromagnetic radiation have very high penetrating powers. For example, some X-rays can pass through different types of human tissues which means that they are very useful in medical imaging (see Figure 9.3.9).

Unfortunately, this useful penetrating property of X-rays comes with inherent dangers. As X-rays pass through a human cell, they can do damage to the tissue, sometimes killing the cells or damaging the DNA in the cell nucleus, leading to harmful cancers. For this reason, a person's exposure to X-rays has to be carefully monitored to avoid harmful side effects.

Similarly, exposure to gamma rays can be very dangerous to human beings. The main natural sources of gamma radiation exposure are the Sun and radioactive isotopes. Fortunately, Earth's atmosphere protects people from most of the Sun's harmful gamma rays, and radioactive isotopes are not commonly found in sufficient quantities to produce harmful doses of radiation.

FIGURE 9.3.9 This X-ray image of a child's hips can be formed because X-rays can pass through human tissue.

9.3 Review

SUMMARY

- Although light exhibits many wave properties, it cannot solely be modelled as a mechanical wave because it can travel through a vacuum.
- Light is a form of electromagnetic radiation.
- Electromagnetic waves are transverse waves made up of mutually perpendicular, oscillating electric and magnetic fields.
- Electromagnetic radiation can be used for a variety of purposes depending on the properties of the waves, which are determined by their frequency.
- Oscillating charges produce electromagnetic waves of the same frequency as the oscillation. Electromagnetic waves cause charges to oscillate at the frequency of the wave.
- Light (that is, all electromagnetic radiation) travels through a vacuum at approximately $c = 3.0 \times 10^8$ m s^{-1}.
- The wave equation $c = f\lambda$ can be used to calculate the frequency and wavelength of electromagnetic waves.

KEY QUESTIONS

1 What is a key difference between light waves and mechanical waves?
- **A** Light waves do not have a measurable wavelength.
- **B** Light waves can travel through a vacuum.
- **C** The speed of light is too fast to be accurately measured.
- **D** Light waves do not undergo diffraction.

2 In an electromagnetic wave, what is the orientation of the changing electric and magnetic fields?
- **A** at 45° to each other
- **B** parallel to each other
- **C** parallel but in opposite directions
- **D** perpendicular to each other

3 What type of electromagnetic radiation would have a wavelength of 200 nm?
- **A** radio waves
- **B** microwaves
- **C** visible light
- **D** ultraviolet light

4 Arrange the types of electromagnetic radiation below in order of increasing wavelength.
FM radio waves / visible light / infrared radiation / X-rays

5 Calculate the frequencies of the following wavelengths of light.
- **a** red of wavelength 656 nm
- **b** yellow of wavelength 589 nm
- **c** blue of wavelength 486 nm
- **d** violet of wavelength 397 nm

6 Although the currently accepted value for the speed of light is 299 792 458 m s^{-1}, this is often approximated as $c = 3.00 \times 10^8$ m s^{-1}. Calculate the percentage error introduced by this approximation.

7 Calculate the wavelength (in nm) of light with a frequency of 6.0×10^{14} Hz.

8 Calculate the wavelength of a UHF (ultra-high frequency) television signal with a frequency of 7.0×10^7 Hz.

9 Calculate the frequency of an X-ray with a wavelength of 200 pm.

10 Calculate the wavelength of the electromagnetic waves produced by a microwave oven.

Chapter review

KEY TERMS

coherent
constructive interference
critical angle
destructive interference
diffraction
diffraction pattern
dispersion
electromagnetic radiation
electromagnetic spectrum
interference
monochromatic
path difference
polarisation
refraction
refractive index
Snell's law
total internal reflection

1 What phenomenon does the diagram below demonstrate?
A diffraction
B interference
C reflection
D refraction

2 Explain how the width of a double-slit interference pattern would change if all the variables were constant but a blue laser was replaced with a green laser.

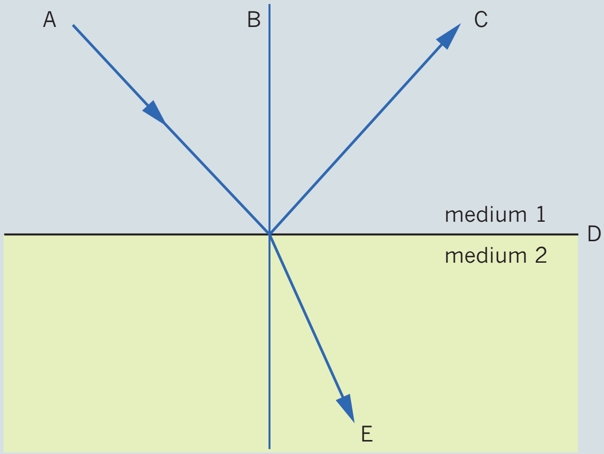
3 Polarisation is an important phenomenon. What does it show about light?
A It can travel instantaneously at an infinite speed.
B It travels faster in materials like water and air than in a vacuum.
C It is a longitudinal wave.
D It is a transverse wave.

4 Explain briefly why snowboarders and sailors are likely to wear polarising sunglasses.

5 Red light (4.5×10^{14} Hz) has a wavelength of 500 nm in water. Calculate the speed of red light in water.

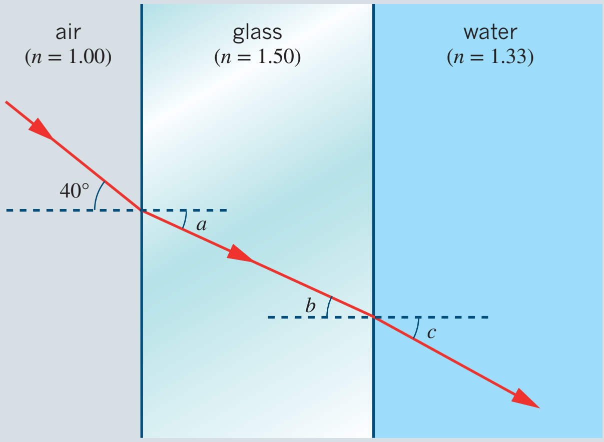
6 Choose the correct answers from those given in bold to complete the following sentence about refraction.
As light travels from quartz (n = 1.46) to water (n = 1.33), its speed **increases**/**decreases** which causes it to refract **away from**/**towards** the normal.

7 The figure represents a situation involving the refraction of light. Identify the correct label for each of the lines from the choices provided: boundary between media, reflected ray, incident ray, normal, refracted ray.



8 The speed of light in air is 3.00×10^8 m s^{-1}.
As light strikes an air–perspex boundary, the angle of incidence is 43.0° and the angle of refraction is 28.5°. Calculate the speed of light in perspex.

9 A ray of light travels from air, through a layer of glass and then into water as shown. Calculate angles *a*, *b* and *c*.



Chapter review *continued*

10 A ray of light exits a glass block. On striking the inside wall of the glass block, the ray makes an angle of 58.0° with the glass-air boundary. The index of refraction of the glass is 1.52. Calculate the:

a angle of incidence

b angle of refraction of the transmitted ray (assuming n_{air} = 1.00)

c angle of deviation

d speed of light in the glass.

11 A narrow beam of white light enters a crown glass prism with an angle of incidence of 30.0°. In the prism, the different colours of light are slowed to varying degrees. The refractive index for red light in crown glass is 1.50 and for violet light the refractive index is 1.53. Calculate the:

a angle of refraction for the red light

b angle of refraction for the violet light

c angle through which the spectrum is dispersed

d speed of the violet light in the crown glass.

12 Calculate the critical angle for light travelling between the following media.

	Incident medium	Refracting medium
a	ice (n = 1.31)	air (n = 1.00)
b	salt (n = 1.54)	air (n = 1.00)
c	cubic zirconia (n = 2.16)	air (n = 1.00)

13 When a light ray refracts, the difference between the angle of incidence and angle of refraction is known as the *angle of deviation*. Sort the following boundaries between media in order of increasing angle of deviation.

A water (n = 1.33) to diamond (n = 2.42)

B water (n = 1.33) to air (n = 1.00)

C air (n = 1.00) to diamond (n = 2.42)

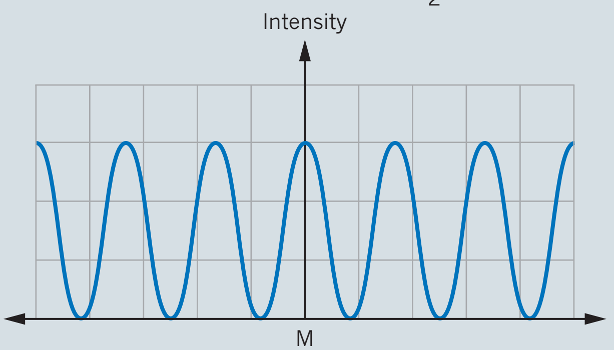
D glass (n = 1.50) to air (n = 1.00)

14 Light of an unknown wavelength emitted by a laser is directed through a pair of thin slits separated by 75 μm. The slits are 4.0 m from a screen on which bright fringes are 3.1 cm apart.

a Calculate the wavelength of the laser light in nm.

b Identify the unknown colour emitted by the laser.

15 The following diagram shows the resulting intensity pattern (simplified) after light from two slits reaches the screen in a double-slit experiment. Copy the diagram into your workbook and circle the points at which the path difference is equal to $1\frac{1}{2}\lambda$.



16 Arrange the types of electromagnetic radiation below in order of decreasing wavelength.
gamma rays, visible, microwaves, radio waves, X-rays, infrared, ultraviolet

17 What form of electromagnetic radiation is used in the following applications?

a mobile phone communication

b night-vision goggles

c medical imaging

18 An AM radio station has a frequency of 612 kHz. If the speed of light is 3×10^8 m s^{-1}, calculate the wavelength of these waves to the nearest metre.

19 Describe Young's experiment and explain why it is considered evidence for the wave theory of light.

20 Explain briefly why a microwave oven is tuned to produce electromagnetic waves of a particular frequency.

CHAPTER 10

Light and matter

How incredible it would be, if it were possible, to put the giants of physics from throughout history—Galileo, Newton, Maxwell, Heisenberg, Bohr, de Broglie, Einstein and others—together in one room for an hour. The most likely outcome is that the hour would be spent in heated debate and discussion. Let's imagine that just one question is posed to them: 'What is light?' None of them would give the same answer as another. Each would have understandings linked to their era. And if another seemingly simple question could be posed—'What is matter?'—the debate could be just as heated!

This chapter will develop your understanding of light and the models used to describe it. It will also introduce the idea that light and matter have more in common than you may have thought.

Key knowledge

By the end of this chapter you will have studied the physics of light and matter, and will be able to:

- analyse the photoelectric effect with reference to:
  - evidence for the particle-like nature of light
  - experimental data in the form of graphs of photocurrent versus electrode potential, and of kinetic energy of electrons versus frequency
  - kinetic energy of emitted photoelectrons: $E_{k\ max} = hf - \phi$, using energy units of joule and electron-volt
  - effects of intensity of incident irradiation on the emission of photoelectrons
- describe the limitation of the wave model of light in explaining experimental results related to the photoelectric effect.
- interpret electron diffraction patterns as evidence for the wave-like nature of matter
- distinguish between the diffraction patterns produced by photons and electrons
- calculate the de Broglie wavelength of matter: $\lambda = \frac{h}{p}$
- compare the momentum of photons and of matter of the same wavelength including calculations using: $p = \frac{h}{\lambda}$
- explain the production of atomic absorption and emission line spectra, including those from metal vapour lamps
- interpret spectra and calculate the energy of absorbed or emitted photons: $\Delta E = hf$
- analyse the absorption of photons by atoms, with reference to:
  - the change in energy levels of the atom due to electrons changing state
  - the frequency and wavelength of emitted photons: $E = hf = \frac{hc}{\lambda}$
- describe the quantised states of the atom with reference to electrons forming standing waves, and explain this as evidence for the dual nature of matter
- interpret the single photon/electron double slit experiment as evidence for the dual nature of light/matter
- explain how diffraction from a single slit experiment can be used to illustrate Heisenberg's uncertainty principle
- explain why classical laws of physics are not appropriate to model motion at very small scales
- compare the production of light in lasers, synchrotrons, LEDs and incandescent lights.

10.1 The photoelectric effect and the dual nature of light

FIGURE 10.1.1 Albert Einstein helped revolutionise our understanding of the nature of light.

At the turn of the 20th century, a number of scientists turned their attention to light phenomena that could not be readily explained using Maxwell's electromagnetic wave model. The study of these phenomena required the development of much more sophisticated models for light, and eventually led to a revolution in the scientific understanding of the nature of energy and matter. One of the scientists who made a significant contribution to this new way of understanding light was Albert Einstein, shown in Figure 10.1.1.

PLANCK'S EQUATION

In 1900, the German physicist Max Planck (see Figure 10.1.2) was studying the spectrum for light emitted by hot objects. Planck and other scientists had discovered that certain features of this spectrum could not be explained using a wave model for light.

Planck proposed a controversial solution to this problem by assuming that light was emitted as discrete packets. He called the discrete packets of energy quanta, and developed an equation for the energy, E, of each **quantum**:

$$E = hf$$

where E is the energy of a quantum of light (J)

f is the frequency of the electromagnetic radiation (Hz)

h is the constant 6.63×10^{-34} J s, now known as Planck's constant

Since electromagnetic radiation is more commonly described according to its wavelength, scientists often combine Planck's equation with the wave equation for light $c = f\lambda$ as follows:

$E = hf$ and $f = \frac{c}{\lambda}$

So

$$E = \frac{hc}{\lambda}$$

At the time, most scientists disregarded Planck's work because the particle model it suggested was so much at odds with the wave model that had become widely accepted as the correct explanation for light.

PHYSICSFILE

Max Karl Ernst Ludwig Planck

Max Planck was a German physicist. At the age of 21 he obtained a PhD in physics, and in 1889 was appointed professor at the university in Berlin. Planck was an author of numerous works about physics—about quantum theory in particular. On 14 December 1900 he presented the correct version of the Wien formula and introduced a new constant—Planck's constant. This date is now recognised as the beginning of the era of quantum mechanics. In 1918, Planck was awarded the Nobel Prize in Physics for the discovery of quantum energy.

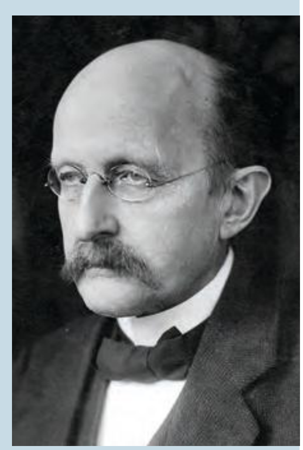


FIGURE 10.1.2 Max Planck (1858–1947).

Worked example 10.1.1

USING PLANCK'S EQUATION

Calculate the energy in joules of a quantum of ultraviolet light that has a frequency of 2.00×10^{15} Hz.	
Thinking	**Working**
Recall Planck's equation.	$E = hf$
Substitute in the appropriate values to solve.	$E = 6.63 \times 10^{-34} \times 2.00 \times 10^{15}$ $= 1.33 \times 10^{-18}$ J

Worked example: Try yourself 10.1.1

USING PLANCK'S EQUATION

Calculate the energy in joules of a quantum of infrared radiation that has a frequency of 3.6×10^{14} Hz.

THE ELECTRON-VOLT

When studying light, the quantities of energy considered are usually very small. These are often so small that the joule is no longer a convenient unit to use. Scientists have therefore adopted a unit called an **electron-volt** (eV). You may have come across it in some optional material in Unit 2. An electron-volt is the amount of energy an electron gains when it moves through a potential difference of 1 V. Since the charge on an electron is -1.6×10^{-19} C, then:

$$1 \text{ eV} = 1e \times 1 \text{ V}$$
$$= 1.6 \times 10^{-19} \text{ C} \times 1 \text{ J C}^{-1}$$
$$= 1.6 \times 10^{-19} \text{ J}$$

Here is a simple way to convert between the units for energy:

To convert a value expressed in J into eV, divide it by 1.6×10^{-19} J eV^{-1}

To convert a value expressed in eV into J, multiply it by 1.6×10^{-19} J eV^{-1}

Worked example 10.1.2

CONVERTING TO ELECTRON-VOLTS

A quantum of light has 1.33×10^{-18} J of energy. Convert this energy to electron-volts.	
Thinking	**Working**
Recall the conversion rate for joules to electron-volts.	1 eV = 1.6×10^{-19} J
Divide the value expressed in joules by 1.6×10^{-19} J eV^{-1} to convert to electron-volts.	$1.33 \times 10^{-18} \div (1.6 \times 10^{-19})$ = 8.3 eV

Worked example: Try yourself 10.1.2

CONVERTING TO ELECTRON-VOLTS

A quantum of light has 2.4×10^{-19} J of energy. Convert this energy to electron-volts.

As seen from worked examples 10.1.1. and 10.1.2, it is easier to compare the relative energies of quanta when they are expressed in eV.

For convenience, Planck's constant can also be given in terms of electron-volts, i.e. $h = 6.63 \times 10^{-34}$ J s

$$= \frac{6.63 \times 10^{-34}}{1.6 \times 10^{-19}}$$

$$= 4.14 \times 10^{-15} \text{ eV s}$$

Worked example 10.1.3

CALCULATING QUANTUM ENERGIES IN ELECTRON-VOLTS

Calculate the energy (in eV) of a quantum of ultraviolet light that has a frequency of 2.0×10^{15} Hz. Use $h = 4.14 \times 10^{-15}$ eV s.	
Thinking	**Working**
Recall Planck's equation.	$E = hf$
Substitute in the appropriate values and solve for E.	$E = 4.14 \times 10^{-15} \times 2.0 \times 10^{15}$ $= 8.3$ eV

Worked example: Try yourself 10.1.3

CALCULATING QUANTUM ENERGIES IN ELECTRON-VOLTS

Calculate the energy (in eV) of a quantum of infrared radiation that has a frequency of 3.6×10^{14} Hz. Use $h = 4.14 \times 10^{-15}$ eV s.

THE PHOTOELECTRIC EFFECT

At the start of the 20th century, another phenomenon that could not be explained using the wave model for light was being observed.

Scientists noticed that when some types of electromagnetic radiation are incident on a piece of metal, the metal becomes positively charged. This positive charge is due to electrons being ejected from the surface of the metal. The electrons became known as **photoelectrons** because they were released due to light or other forms of electromagnetic radiation. The phenomenon is known as the **photoelectric effect**.

A common apparatus used to observe the photoelectric effect consists of a clean metal surface (the cathode), illuminated with light from an external source. If the light causes photoelectrons to be emitted, they are detected at the anode. The flow of electrons is called the **photocurrent** and is registered by a sensitive ammeter.

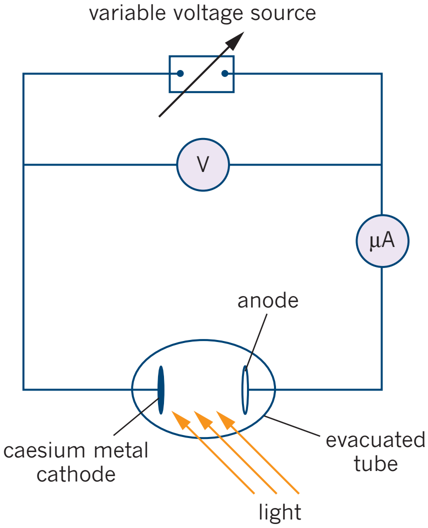


FIGURE 10.1.3 Circuit diagram of an experimental investigation of the photoelectric effect.

A typical circuit used to investigate the photoelectric effect (see Figure 10.1.3) includes a variable voltage supply, which can be used to make the cathode negative (and the anode positive). When this is done, the photoelectrons will be helped by the resulting electric field across the gap to the anode. This happens because the electrons will be repelled by the negative potential at the cathode and attracted to the positive potential at the anode. As a result, a maximum possible current will be measured. Alternatively, the voltage may be adjusted to make the cathode positive and the anode negative. This repels the photoelectrons and slows them down. As the anode voltage is increased, the photoelectrons are repelled more and more until the photocurrent drops to zero.

Using the apparatus shown in Figure 10.1.3, the German physicist Philipp Lenard made a number of surprising discoveries about the photoelectric effect. He won the Nobel Prize in Physics in 1905 for his discoveries.

Lenard used a filter to vary the frequency of the incident light. He discovered that, for a particular cathode metal, there is a certain frequency below which no photoelectrons are observed. This is called the **threshold frequency**, f_0. For frequencies of light greater than the threshold frequency (i.e. $f > f_0$), photoelectrons will be collected at the anode and registered as a photocurrent. For frequencies below the threshold frequency (i.e. $f < f_0$), no photoelectrons will be detected.

Lenard also discovered that, for light that has a frequency greater than the threshold frequency, i.e. $f > f_0$, the rate at which the photoelectrons are produced varies in proportion with the intensity (brightness) of the incident light as shown in Figure 10.1.4.

This graph shows a number of important properties of the photoelectric effect.

- When the light intensity increases, the photocurrent increases.
- When the applied voltage is positive, photoelectrons are attracted to the collector electrode (anode). A small positive voltage is enough to ensure that every available photoelectron is collected. The current therefore reaches a maximum value and remains there even if the voltage is increased.
- When the applied voltage is negative, photoelectrons are attracted back towards the illuminated cathode and repelled by the collector electrode (anode) and the photocurrent is reduced. The photocurrent is reduced because fewer and fewer photoelectrons have the energy to overcome the opposing electric potential. There is a voltage, V_0, for which no photoelectrons reach the collector. This is known as the **stopping voltage**. For a particular frequency of light on a particular metal, this stopping voltage is a constant.

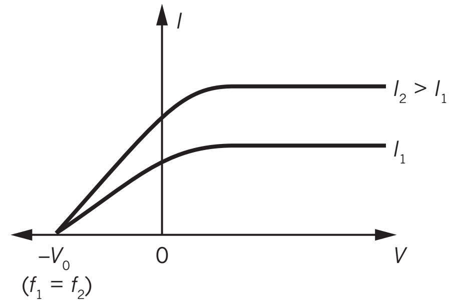


FIGURE 10.1.4 Photocurrent (I) plotted as a function of the voltage (V) applied between the cathode and the anode for different light intensities. For brighter light ($I_2 > I_1$) of the same frequency ($f_1 = f_2$), there is a higher photocurrent, but the same stopping voltage, V_0.

Recall from earlier studies of electricity that the work done on a charge (by an applied voltage) is given by $W = qV$. In this case, the voltage used is designated the stopping voltage, V_0, and the charge value is equal to the magnitude of the charge on an electron, q_e, 1.6×10^{-19} C. Hence the work done on the electron is given by $W = q_e V_0$. Since the stopping voltage is large enough to stop even the fastest-moving electrons from reaching the anode, this expression gives the value of the maximum possible kinetic energy of the emitted photoelectrons. For example, should the stopping voltage be 2.5 V, then the maximum kinetic energy of any photoelectron is 2.5 eV.

When light sources of the same intensity but different *frequencies* are used, they produce the same maximum current. However, the higher frequency light has a higher stopping voltage (see Figure 10.1.5).

Finally, as long as the incident light has a frequency above the threshold frequency of the cathode material, photoelectrons are found to be emitted without any appreciable time delay. This fact holds true regardless of the intensity of the light.

When illuminated with light above the threshold frequency, some photoelectrons are emitted from the first layer of atoms at the surface of the metal and have the maximum kinetic energy possible. Other photoelectrons come from deeper inside the metal and lose some of their kinetic energy due to collisions on their way to the surface. Hence, the emitted photoelectrons have a range of kinetic energies from the maximum value downwards.

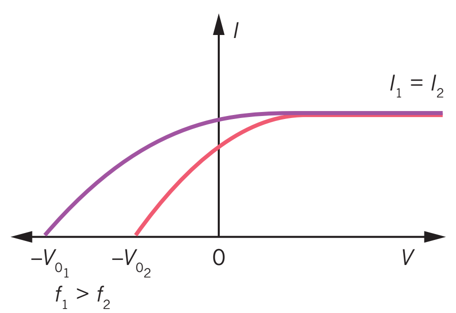


FIGURE 10.1.5 Photocurrent (I) plotted as a function of the voltage (V) applied between the cathode and the anode for different frequencies ($f_1 > f_2$) of incident light with the same intensity ($I_1 = I_2$). Both frequencies produce the same maximum photocurrent; however, light with the higher frequency requires a larger stopping voltage.

EXPLAINING THE PHOTOELECTRIC EFFECT

The characteristics of the photoelectric effect could not be explained using a wave model of light. According to the wave model, the frequency of light should be irrelevant to whether or not photoelectrons are ejected. Since a wave is a form of continuous energy transfer, it would be expected that even low-frequency light should transfer enough energy to emit photoelectrons if left incident on the metal for long enough. Similarly, the wave model predicts that there should be a time delay between the light striking the metal and photoelectrons being emitted, as the energy from the wave builds up in the metal over time.

The dual nature of light

In 1905, Albert Einstein proposed a solution to this problem. Einstein drew on Planck's earlier work by assuming that light exists as particles, or **photons** (like Planck's 'quanta'), each with an energy of $E = hf$. This assumption made the properties of the photoelectric effect relatively easy to explain.

Einstein's work was actually a significant extension of Planck's ideas. Although Planck had assumed that light was being emitted in quantised packets, he never questioned the assumption that light was fundamentally a wave phenomenon.

Einstein's work went further, challenging scientists' understanding of the nature of light itself.

Einstein and the photoelectric effect

Einstein identified that, for a particular metal, the amount of energy required to eject a photoelectron is a constant value that depends on the strength of the bonding within the metal. This energy was called the **work function**, ϕ, of the metal. For example, the work function of lead is 4.14 eV, which means that 4.14 eV of energy is needed to just release one electron from the surface of a piece of lead.

According to Einstein's model, shining light on the surface of a piece of metal is equivalent to bombarding it with photons. When a photon strikes the metal, it can transfer its energy to an electron. That is, a single photon can interact with a single electron, transferring all of its energy at once to the electron. What happens next depends on whether or not the photon contains enough energy to overcome the work function.

If the energy of the photon is less than the work function, then photoelectrons will not be released as the electrons will not gain enough energy to let them break free of the lead atoms. For example, the photons of violet light ($f = 7.50 \times 10^{14}$ Hz) each contain 3.11 eV of energy.

$$E = hf$$
$$= 4.14 \times 10^{-15} \times 7.50 \times 10^{14}$$
$$= 3.11 \text{ eV}$$

This means that violet light shining on lead would not release photoelectrons since the energy of each photon, 3.11 eV, is less than the work function of lead, 4.14 eV.

However, ultraviolet photons of frequency 1.20×10^{15} Hz each contain 4.97 eV of energy.

$$E = hf$$
$$= 4.14 \times 10^{-15} \times 1.20 \times 10^{15}$$
$$= 4.97 \text{ eV}$$

Therefore ultraviolet light of this frequency would release photoelectrons from the lead since the energy of each photon, 4.97 eV, is greater than the work function of lead, 4.14 eV.

Each metal has a threshold frequency—this is the frequency at which the photons have an energy equal to the work function of the metal:

$$\phi = hf_0$$

where ϕ is the work function (J or eV)

h is Planck's constant

f_0 is the threshold frequency for that metal (Hz)

Worked example 10.1.4

CALCULATING THE WORK FUNCTION OF A METAL

Calculate the work function (in J and eV) for aluminium, which has a threshold frequency of 9.8×10^{14} Hz.

Thinking	Working
Recall the formula for work function.	$\phi = hf_0$
Substitute the threshold frequency of the metal into this equation.	$\phi = 6.63 \times 10^{-34} \times 9.8 \times 10^{14}$ $= 6.5 \times 10^{-19}$ J
Convert this energy from J to eV.	$\phi = \frac{6.5 \times 10^{-19}}{1.6 \times 10^{-19}}$ $= 4.1$ eV

Worked example: Try yourself 10.1.4

CALCULATING THE WORK FUNCTION OF A METAL

Calculate the work function (in J and eV) for gold, which has a threshold frequency of 1.2×10^{15} Hz.

THE KINETIC ENERGY OF PHOTOELECTRONS

If the energy of the photon is greater than the work function of the metal, then a photoelectron is released. The remainder of the energy in excess of the work function is transformed into the kinetic energy of the photoelectron.

Einstein described this relationship with his photoelectric equation:

$$E_{k\,max} = hf - \phi$$

where $E_{k\,max}$ is the maximum kinetic energy of an emitted photoelectron (J or eV)

ϕ is the work function of the metal (J or eV)

h is Planck's constant

f is the frequency of the incident photon (Hz)

Graphing Einstein's equation results in a linear (straight line) graph like the one shown in Figure 10.1.6. Such a graph is useful because it clearly shows key information such as the work function and threshold frequency for a particular metal.

Einstein's equation, $E_{k\,max} = hf - \phi$ can be compared with the equation of a straight line, $y = mx + c$. In making this comparison, it can be seen that extrapolating (extending) the graph back to the vertical axis will give the magnitude of the work function, ϕ (see Figure 10.1.6). The gradient of the graph is Planck's constant, h. From the graph it is also apparent how, as soon as the threshold frequency is exceeded, an electron is able to be ejected and escape with some kinetic energy. The greater the frequency of the light, the greater the kinetic energy of the photoelectron. At the threshold frequency, electrons are no longer bound to the metal, but they have no kinetic energy.

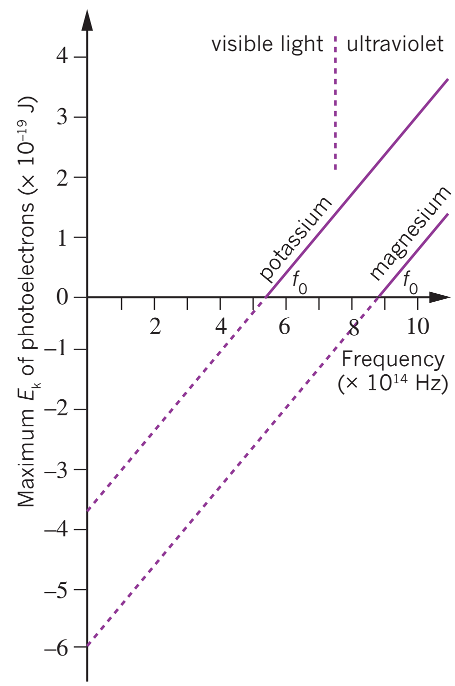


FIGURE 10.1.6 Magnesium has a high threshold frequency, which is in the ultraviolet region. The threshold frequency for potassium is in the visible region. The gradient of the graph for each metal is Planck's constant, h. The x-intercept gives the threshold frequency, f_0. The magnitude of the y-intercept gives the work function, ϕ.

Worked example 10.1.5

CALCULATING THE KINETIC ENERGY OF PHOTOELECTRONS

Calculate the kinetic energy (in eV) of the photoelectrons emitted from lead by ultraviolet light with a frequency of 1.2×10^{15} Hz. The work function of lead is 4.14 eV. Use $h = 4.14 \times 10^{-15}$ eV s.

Thinking	Working
Recall Einstein's photoelectric equation.	$E_{k\,max} = hf - \phi$
Substitute values into this equation.	$E_{k\,max} = 4.14 \times 10^{-15} \times 1.2 \times 10^{15} - 4.14$ $= 4.97 - 4.14$ $= 0.83$ eV

Worked example: Try yourself 10.1.5

CALCULATING THE KINETIC ENERGY OF PHOTOELECTRONS

Calculate the kinetic energy (in eV) of the photoelectrons emitted from lead by ultraviolet light with a frequency of 1.5×10^{15} Hz. The work function of lead is 4.14 eV. Use $h = 4.14 \times 10^{-15}$ eV s.

PHYSICSFILE

Albert Einstein

Although Albert Einstein is most famous for his work on relativity (and its related equation $E = mc^2$), he gained his Nobel Prize 'for his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect'. His work on relativity was never formally recognised with a Nobel Prize.

Resistance to the quantum model of light

This new particle or 'quantum' model of light was not initially well received by the scientific community. It had already been well established that a discrete, particle model for light could not explain many of light's properties such as polarisation and the interference patterns produced in Young's experiment.

Most scientists believed instead that wave explanations for the photoelectric effect would eventually be found. However, eventually the quantum model of light was accepted and the Nobel Prize in Physics was awarded to both Planck (1918) and Einstein (1921) for their ground-breaking work in this field.

PHYSICS IN ACTION

Photovoltaic cells

The photovoltaic cells that are used in many solar panels work on the principle of the photoelectric effect (see Figure 10.1.7). Sunlight falling on the solar panel provides energy that causes photoelectrons to be emitted as a current that can be used to drive electrical appliances.

However, whereas many photoelectric-effect experiments use high-energy photons of ultraviolet light, photovoltaic cells use materials that will produce photoelectrons when exposed to visible light. Most commonly, these are semi-conducting materials based on silicon 'doped' with small amounts of other elements.

Although solar cells are designed to produce the highest current possible from sunlight, most commercially available solar cells have an energy efficiency of less than 20%. Scientists hope to improve this in order to make solar cells an economic alternative to fossil fuels for large-scale energy generation.

FIGURE 10.1.7 Solar panels are used to convert sunlight into electrical energy using the photoelectric effect.

10.1 Review

SUMMARY

- On the atomic level, electromagnetic radiation is emitted or absorbed in discrete packets or quanta called photons.
- A number of phenomena related to the behaviour of light, such as the photoelectric effect, can only be explained using the concept of photons, or light quanta.
- The energy of a photon is proportional to its frequency: $E = hf = \frac{hc}{\lambda}$
- The constant of proportionality, Planck's constant, can be determined experimentally using the photoelectric effect.
- The electron-volt is an alternative (non-SI) unit of energy: 1 eV = 1.6×10^{-19} J.
- The photoelectric effect is the emission of photoelectrons from a clean metal surface due to incident light whose frequency is greater than a threshold frequency, f_0.
  - If $f < f_0$, no electrons are released.
  - If $f > f_0$, the *rate* of electron release (the photocurrent) is proportional to the intensity of the light and occurs without any time delay.
- Increasing the forward voltage does not alter the rate of electron release (i.e. the photocurrent).
- By experiment, the maximum kinetic energy for the electrons $E_{k\,max}$ (i.e. that of the fastest electron) can be found by using a reverse voltage called the stopping voltage, V_0.
- $E_{k\,max} = q_eV_0$, where q_e is the charge on an electron.
- The work function, ϕ, for the metal is given by $\phi = hf_0$, and is different for each metal. If the frequency of the incident light is greater than the threshold frequency, then a photoelectron will be ejected with some kinetic energy up to a maximum value.
- A graph of $E_{k\,max}$ versus frequency will have a gradient equal to Planck's constant, h, and a 'y intercept' equal to the work function, ϕ.
- The maximum kinetic energy of the photoelectrons emitted from a metal is the energy of the photons minus the work function, ϕ, of the metal: $E_{k\,max} = hf - \phi$.
- The wave approach to light could not explain various features of the photoelectric effect: the existence of a threshold frequency, the absence of a time delay when using very weak light sources, and increased intensity of light resulting in a greater rate of electron release rather than increased electron energy.
- Einstein used Planck's concept of a photon to explain the photoelectric effect, stating that each electron release was due to an interaction with only one photon.
- The photon model of light explained the existence of a threshold frequency for each metal, the absence of a time delay for the photocurrent even for weak light sources and why brighter light resulted in a higher photocurrent.

KEY QUESTIONS

1 Calculate the energies (in joules) of the following wavelengths of light:

	Colour	Wavelength (nm)
a	red	656
b	yellow	589
c	blue	486
d	violet	397

2 When light shines on a metal surface, why might the metal become positively charged?

3 Which of the following statements about the photoelectric effect are true and which are false? For those that are false, rewrite them to make them correct.

a When the intensity of light shining on the surface of the metal increases, the photocurrent increases.

b When light sources of the same intensity but different frequencies are used, the higher frequency light has a higher stopping voltage and produces a higher maximum current than the lower frequency.

c When the applied voltage is positive, photoelectrons are attracted to the collector electrode.

4 Calculate the work functions (in electron-volts) of the following metals:

	Metal	Threshold frequency ($\times 10^{15}$ Hz)
a	lead	1.0
b	iron	1.1
c	platinum	1.5

10.1 Review *continued*

5 In an experiment on the photoelectric effect, different frequencies of light were shone on a piece of magnesium with a work function of 3.66 eV. Identify which of the frequencies listed would be expected to produce photoelectrons.

A 3.0×10^{14} Hz
B 5.0×10^{14} Hz
C 7.0×10^{14} Hz
D 9.0×10^{14} Hz

6 Light with a frequency of 9.0×10^{14} Hz is shone onto a piece of magnesium with a work function of 3.66 eV. Calculate the maximum kinetic energy, in electron-volts, of the emitted photoelectrons.

7 Blue light with a wavelength of 475 nm is shone on a piece of sodium with a work function of 2.36 eV. Calculate the maximum kinetic energy, in electron-volts, of the emitted photoelectrons.

8 The metal sodium has a work function of 1.81 eV. Which of the following types of electromagnetic radiation would cause photoelectrons to be emitted?

A infrared radiation, $\lambda = 800$ nm
B red light, $\lambda = 700$ nm
C violet light, $\lambda = 400$ nm
D ultraviolet radiation, $\lambda = 300$ nm

9 Which of the following statements are true and which are false with respect to the value of the stopping voltage obtained when light is incident on a metal cathode? For those that are false, rewrite them to make them true.

a The stopping voltage indicates how much work must be done to stop the most energetic photoelectrons.
b The stopping voltage is reached when the photocurrent is reduced almost to zero.
c If only the intensity of the incident light is increased, the stopping voltage will not alter.
d For a given metal, the value of the stopping voltage is affected only by the frequency of the incident light.

10 Yellow–green light of wavelength 500 nm shines on a metal whose stopping voltage is found to be 0.80 V. Calculate the work function of the metal in electron-volts.

10.2 The quantum nature of light and matter

In order to explain the photoelectric effect, Einstein used the photon concept that Planck had developed. However, like many great discoveries in science, the development of the quantum model of light raised almost as many questions as it answered. It has already been well established that a wave model was needed to explain phenomena such as diffraction and interference. How could these two contradictory models be reconciled to form a comprehensive theory of light?

Answering this question was one of the great scientific achievements of the 20th century and led to the extension of the quantum model to matter as well as energy. It led to a fundamental shift in the way the universe is viewed. Some of the great scientists of that time are shown in the historic photograph in Figure 10.2.1.

FIGURE 10.2.1 This photo shows the 5th Solvay conference in Brussels in 1927, which was attended by great scientists including Albert Einstein, Max Planck, Niels Bohr, Marie Curie, Paul Dirac, Erwin Schrödinger and Louis de Broglie. All of these scientists contributed to the current knowledge of the universe, the atom and quantum mechanics.

WAVE–PARTICLE DUALITY

In many ways, the wave and particle models for light seem fundamentally incompatible. Waves are continuous and are described in terms of wavelength and frequency. Particles are discrete and are described by physical dimensions such as their mass and radius.

In order to understand how these two sets of ideas can be used together, it is important to remember that scientists describe the universe using models. Models are analogies that are used to illustrate certain aspects of reality that might not be immediately apparent.

Physicists have come to accept that light is not easily compared to any other physical phenomenon. In some situations, light has similar properties to a wave; in other situations, light behaves more like a particle. This understanding is called **wave–particle duality** (Figure 10.2.2). Although this may seem somewhat paradoxical and counter-intuitive, in the century since Einstein did his work establishing quantum theory, many experiments have supported this duality and no scientist has (yet) come up with a better explanation.

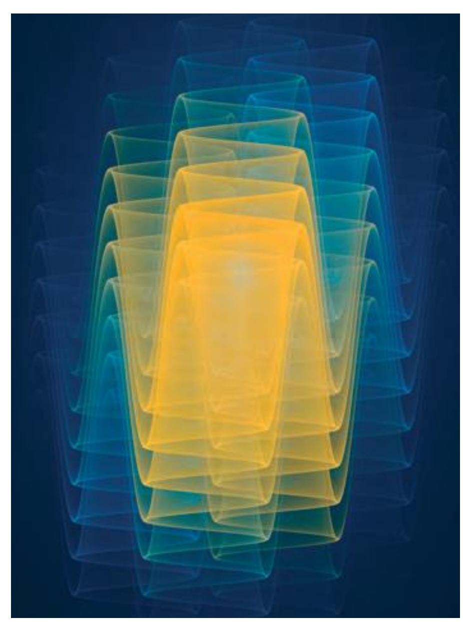


FIGURE 10.2.2 An artist's attempt to represent wave–particle duality.

Experimental evidence for the dual nature of light

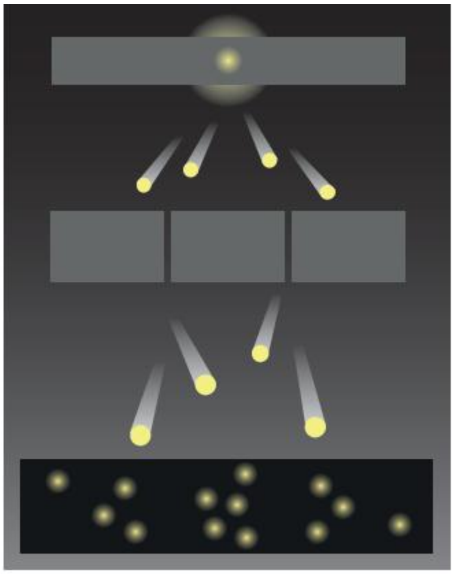


FIGURE 10.2.3 An interference pattern can be built up over time by a series of single photons passing through an apparatus like that used in Young's experiment, demonstrating the wave–particle duality of light.

In the early years of quantum theory, some scientists believed that the wave properties of light observed in Young's double-slit experiment may have been due to some sort of interaction between photons as they passed through the slits together.

To test this, experiments were done with light sources that were so dim that scientists were confident that only one photon was passing through the apparatus at a time. In this way, any interactions between photons could be eliminated. Over time, these experiments produced identical interference patterns to those done with bright sources (see Figure 10.2.3) thus demonstrating the dual nature of light.

Interestingly, when a detector is used to measure which slit the photon passes through, the wave pattern disappears and the photon acts like a particle.

De Broglie's wave–particle theory

In 1924, the French physicist Louis de Broglie proposed a ground-breaking theory. He suggested that since light (which had long been considered to be a wave) sometimes demonstrated particle-like properties, then perhaps matter (which was considered to be made up of particles) might sometimes demonstrate wave-like properties.

He quantified this theory by predicting that the wavelength of a particle would be given by the equation:

$$\lambda = \frac{h}{p}$$

where λ is the wavelength of the particle (m)

p is the momentum of the particle (kg m s^{-1})

h is Planck's constant

This is also commonly written as:

$$\lambda = \frac{h}{mv}$$

where m is the mass of the particle (kg)

v is the velocity of the particle (m s^{-1})

The wavelength that de Broglie described, λ, is referred to as the **de Broglie wavelength** of matter.

PHYSICSFILE

Louis Victor Pierre Raymond de Broglie (1892–1987)

Louis de Broglie (Figure 10.2.4) was a French physicist. In 1924 he wrote a doctoral thesis entitled Recherches sur la théorie des quanta (Research on quantum theory), in which he presented his theory of the wave properties of particles—the de Broglie wave theory, based on the works of Einstein and Planck on wave–particle duality. Later, de Broglie developed his thesis and formulated the final de Broglie hypothesis. In 1929 he was awarded the Nobel Prize for his research. By applying de Broglie's theory it was possible, for example, to construct an electron microscope.

FIGURE 10.2.4 Louis de Broglie

Worked example 10.2.1

CALCULATING THE DE BROGLIE WAVELENGTH

Electrons in a famous experiment known as the Davisson–Germer experiment travelled at about 4.0×10^6 m s^{-1}. Calculate the de Broglie wavelength of these electrons if the mass of an electron is 9.11×10^{-31} kg.	
Thinking	**Working**
Recall de Broglie's equation.	$\lambda = \frac{h}{mv}$
Substitute the appropriate values into the equation and solve it.	$\lambda = \frac{h}{mv}$ $= \frac{6.63 \times 10^{-34}}{9.11 \times 10^{-31} \times 4 \times 10^{6}}$ $= 1.8 \times 10^{-10}$ m or 0.18 nm

Worked example: Try yourself 10.2.1

CALCULATING THE DE BROGLIE WAVELENGTH

Calculate the de Broglie wavelength of a proton travelling at 7.0×10^5 m s^{-1}. The mass of a proton is 1.67×10^{-27} kg.

Worked example 10.2.2

CALCULATING THE DE BROGLIE WAVELENGTH OF A MACROSCOPIC OBJECT

Calculate the wavelength of a cricket ball of mass m = 160 g travelling at 150 km h^{-1}.	
Thinking	**Working**
Convert mass and velocity to SI units.	m = 160 g = 0.16 kg v = 150 ÷ 3.6 = 42 m s^{-1}
Recall de Broglie's equation.	$\lambda = \frac{h}{mv}$
Substitute the appropriate values into the equation and solve it.	$\lambda = \frac{h}{mv}$ $= \frac{6.63 \times 10^{-34}}{0.16 \times 42}$ $= 9.9 \times 10^{-35}$ m

Worked example: Try yourself 10.2.2

CALCULATING THE DE BROGLIE WAVELENGTH OF A MACROSCOPIC OBJECT

Calculate the de Broglie wavelength of a person with m = 66 kg running at 36 km h^{-1}.

It can be seen from worked examples 10.2.1 and 10.2.2 that the wavelength of an electron is smaller than that of visible light, but is still large enough to be measurable. However, the wavelength of an everyday object such as a cricket ball is extremely small (9.9×10^{-35} m). Hence, you will never notice the wave properties of everyday objects. To illustrate this, consider the observable wave behaviour of diffraction. Recall that for diffraction to be noticeable, the size of the wavelength needs to be comparable to the size of the gap or obstacle. Therefore for an everyday object, with its tiny wavelength, to produce a noticeable diffraction, it would need to pass through a gap much smaller than a fraction of a proton diameter!

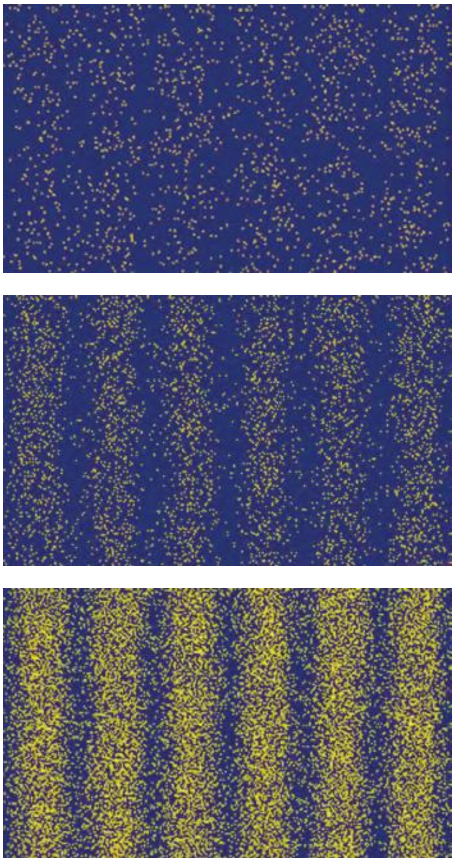


FIGURE 10.2.6 An electron diffraction pattern like the one observed by Davisson and Germer can be built up over time from repeated observations.

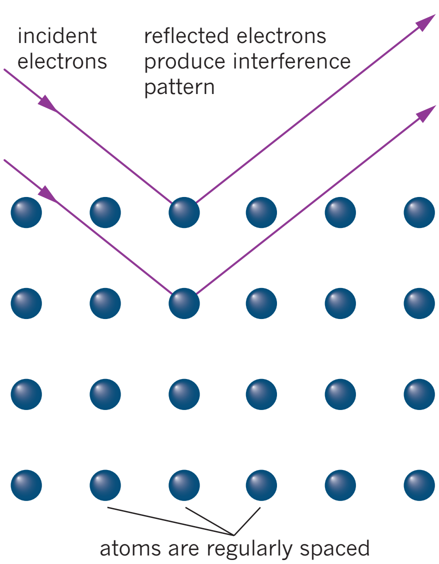


FIGURE 10.2.7 Electrons reflecting from different layers within the crystal structure create an interference pattern like those produced by a diffraction grating.

ELECTRON DIFFRACTION PATTERNS

De Broglie's prediction that matter could exhibit wave-like behaviour was controversial. However, it was experimentally confirmed by the Americans Davisson and Germer in 1927 when they observed diffraction patterns being produced when they bombarded the surface of a piece of nickel with electrons (see Figure 10.2.5).

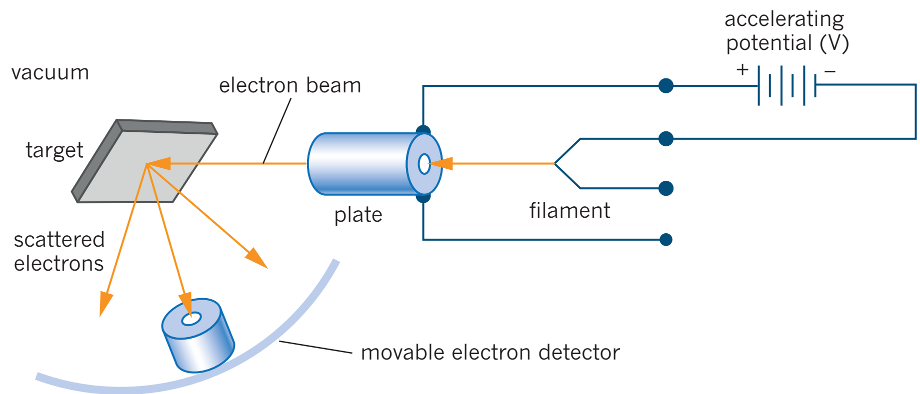


FIGURE 10.2.5 The Davisson and Germer apparatus to show electron scattering.

They used an electron 'gun' which provided a beam of electrons. The speed of the electrons was known because they had been accelerated through a known voltage. The detector could be swung around on an axis so that it could intercept electrons scattered from the nickel target in any direction in the plane shown.

Davisson and Germer found that as they moved their detector through the different scattering angles, they encountered a sequence of maximum and minimum intensities (see Figure 10.2.6).

Clearly, the electrons were being scattered by the different layers within the crystal lattice (see Figure 10.2.7) and were undergoing interference. When Davisson and Germer analysed the diffraction pattern to determine the wavelength of the 'electron waves', they calculated a value of 0.14 nm, which was consistent with de Broglie's hypothesis.

Worked example 10.2.3

WAVELENGTH OF ELECTRONS FROM AN ELECTRON GUN

Find the de Broglie wavelength of an electron that has been accelerated from rest through a potential difference of 75 V. The mass of an electron is 9.11×10^{-31} kg and the magnitude of the charge on an electron is 1.6×10^{-19} C.

Thinking	Working
Calculate the kinetic energy of the electron from the work done on it by the electric potential. Recall from earlier chapters that $W = qV$.	$W = qV$ $= 1.6 \times 10^{-19} \times 75$ $= 1.2 \times 10^{-17}$ J
Calculate the velocity of the electron.	$E_k = \frac{1}{2} mv^2$ $v = \sqrt{\frac{2E_k}{m}}$ $= \sqrt{\frac{2 \times 1.2 \times 10^{-17}}{9.11 \times 10^{-31}}}$ $= 5.1 \times 10^6$ m s^{-1}
Use de Broglie's equation to calculate the wavelength of the electron.	$\lambda = \frac{h}{mv}$ $= \frac{6.63 \times 10^{-34}}{9.11 \times 10^{-31} \times 5.1 \times 10^6}$ $= 1.4 \times 10^{-10}$ m $= 0.14$ nm

Worked example: Try yourself 10.2.3

WAVELENGTH OF ELECTRONS FROM AN ELECTRON GUN

Find the de Broglie wavelength of an electron that has been accelerated from rest through a potential difference of 50 V. The mass of an electron is 9.11×10^{-31} kg and the magnitude of the charge on an electron is 1.6×10^{-19} C.

Comparing the wavelengths of photons and electrons

In the same year that Davisson and Germer conducted their experiment, other supporting evidence came from G. P. Thomson (son of J. J. Thomson, discoverer of the electron). Rather than scatter an electron beam from a crystal, Thomson produced a diffraction pattern by passing a beam of electrons through a tiny crystal. Thomson then repeated his experiment, using X-rays of the same wavelength in place of the electrons. The X-ray diffraction pattern was almost identical to the one made with electrons, as shown in Figure 10.2.8.

As the diffraction patterns obtained for the X-ray photons and electrons were the same, and as both were passed through the same 'gaps' to obtain this diffraction pattern, then an important conclusion could be made. The electrons must have a similar wavelength to the X-rays. Since their wavelengths are similar, the momenta of the electrons and the X-ray photons must also be comparable (but not their speeds).

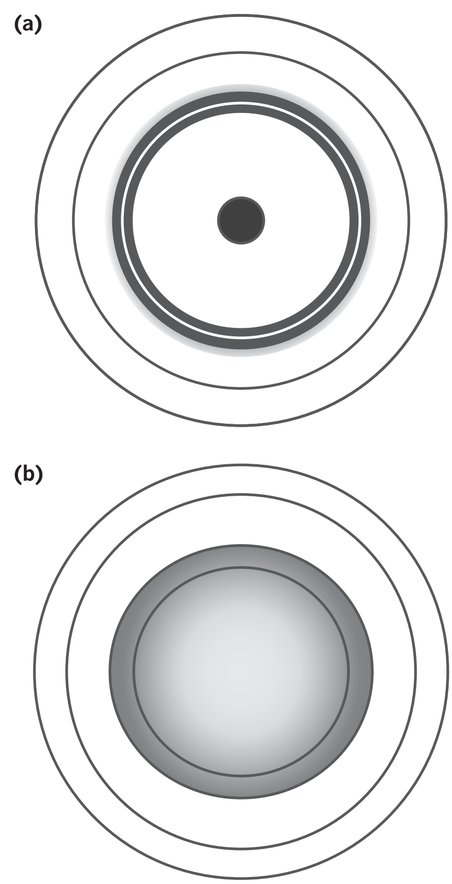


FIGURE 10.2.8 These diffraction patterns were taken by using (a) X-rays and (b) a beam of electrons with the same target crystal. Their similarity suggests a wave-like behaviour for the electrons and a similar electron de Broglie wavelength to that of X-rays.

EXTENSION

Electron microscopes

The discovery of the wave properties of electrons had an important practical application in the invention of the electron microscope. Just as an optical microscope makes use of the wave properties of photons to magnify tiny objects, so too can the wave properties of electrons be used to create magnified images (See Figure 10.2.9).

One of the limitations of an optical microscope is that it can only create a clear image of structures that are similar in size to the wavelength of the light being used. This is because the light diffracts around these structures. So a light microscope is only useful for seeing things down to about 390 nm, the lower wavelength end of the visible light spectrum.

However, the wavelength of a beam of electrons is often smaller than the wavelength of a beam of visible light. This means that electron microscopes can create images with much finer detail than optical microscopes.

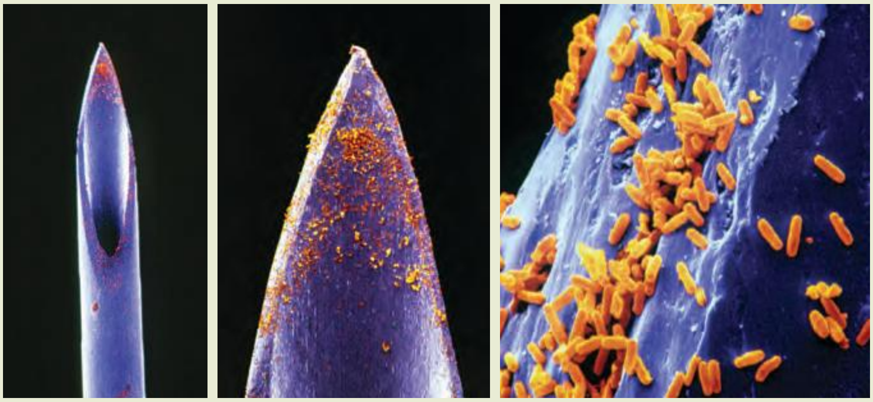


FIGURE 10.2.9 Images formed by an electron microscope: rod-shaped bacteria (orange) clustered on the point of a syringe used to administer injections. The magnifications are (a) ×9, (b) ×36 and (c) ×560 at 35 mm size.

PHOTON MOMENTUM

An interesting corollary of de Broglie's hypothesis, $\lambda = \frac{h}{p}$, is that if a particle like an electron has a wavelength, λ, then a photon must have a momentum, p. This is quite counter-intuitive, since photons do not have any mass and all photons travel at the speed of light, c. Nevertheless, de Broglie's equation allows the momentum of a photon to be calculated.

Worked example 10.2.4

CALCULATING PHOTON MOMENTUM

Calculate the momentum of a photon of red light with a wavelength of 650 nm.	
Thinking	**Working**
Convert 650 nm to m.	650 nm = 650×10^{-9} m
Transpose de Broglie's equation to make momentum the subject.	$\lambda = \frac{h}{p}$ $p = \frac{h}{\lambda}$
Substitute in the appropriate values and solve for p.	$p = \frac{h}{\lambda}$ $= \frac{6.63 \times 10^{-34}}{650 \times 10^{-9}}$ $= 1.02 \times 10^{-27}$ kg m s^{-1}

Worked example: Try yourself 10.2.4

CALCULATING PHOTON MOMENTUM

Calculate the momentum of a photon of blue light with a wavelength of 450 nm.

Clearly, the momentum of a single photon is tiny, which is why you will not feel any physical 'pressure' when light falls on you. However, it is possible to measure 'light pressure' using very sensitive equipment.

PHYSICS IN ACTION

Solar sailing

In interplanetary space, where other forces like friction are negligible, light pressure can actually be used as a form of propulsion. Spacecraft such as the Mariner 10 and MESSENGER spacecraft, which both flew past Mercury and Venus, used deceleration caused by solar pressure to conserve fuel.

More recently, the Japanese Aerospace Exploration Agency launched IKAROS (Interplanetary Kite-craft Accelerated by Radiation Of the Sun). IKAROS (see Figure 10.2.10) is the first spacecraft to draw its primary propulsion from a *solar sail*. A traditional sail propels a ship using the change of momentum that occurs when air molecules bounce off it. Similarly, a solar sail gains propulsion from changes in photon momentum as light is reflected from it. The IKAROS spacecraft has a 196 m^2 reflective sail which produces a thrust of 1.12 mN.

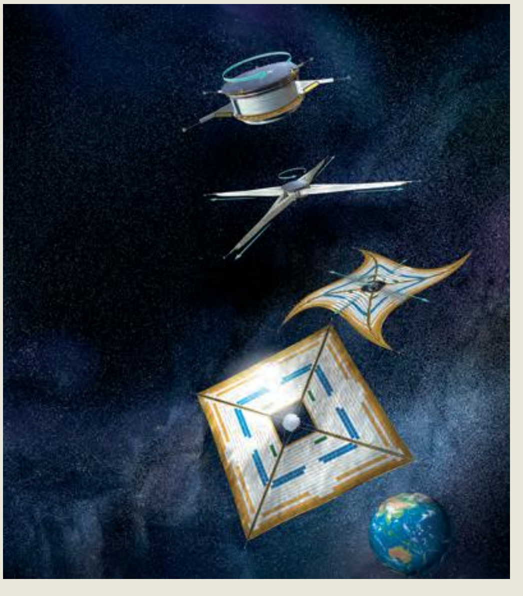


FIGURE 10.2.10 The IKAROS spacecraft is the first interplanetary spacecraft to use solar-sail technology.

10.2 Review

SUMMARY

- On the atomic level, energy and matter exhibit the characteristics of both waves and particles.
- The wavelength of a particle is given by the de Broglie equation:

  $\lambda = \frac{h}{p}$

  i.e. $\lambda = \frac{h}{mv}$
- Young's double-slit experiment is explained with a wave model but produces the same interference and diffraction patterns when one photon at a time or one electron at a time is passed through the slits.
- In particle-scattering experiments, beams of particles (electrons usually) are made to travel with a speed so that their matter wavelength approximates the interatomic spacing in a crystal. Consequently, a diffraction pattern is produced which can only be explained if matter has a wave-like nature.
- If photons and matter particles being scattered by the same crystal sample produce the same fringe spacing, then they must have the same wavelength and momentum.
- All matter, like light, has a dual nature. Through everyday experience matter is particle-like, but under some situations it has a wave-like nature. This symmetry in nature—the dual nature of light and matter—is referred to as wave–particle duality.

KEY QUESTIONS

1 What is the de Broglie wavelength of an electron travelling at 1.0×10^{6} m s^{-1}?

2 Calculate the speed of an electron that has a de Broglie wavelength of 4.0×10^{-9} m.

3 Which of the following conclusions can be drawn from Louis de Broglie's investigation into the existence of matter waves?

A all particles exhibit wave behaviour

B only moving particles exhibit wave behaviour

C only charged particles exhibit wave behaviour

D only moving, charged particles exhibit wave behaviour

4 In an experiment to determine the structure of a crystal, identical diffraction patterns were formed by a beam of electrons and a beam of X-rays with a frequency of 8.6×10^{18} Hz.

a Calculate the wavelength of the electrons.

b Calculate the speed of the electrons.

5 Explain why a cricket player does not have to consider the wave properties of a cricket ball while batting.

6 Explain why it is impossible for individual atoms to be observed by an electron microscope.

7 At what speed would a proton be travelling if it were to have the same wavelength as a gamma ray of energy 6.63×10^{-14} J? (Mass of a proton = 1.67×10^{-27} kg.)

8 A charge q of mass m is accelerated from rest through a potential difference of V. Derive an expression that defines the de Broglie wavelength of the mass, λ, in terms of q, m and V.

9 A corollary of de Broglie's work on matter waves is that photons can be considered to have momentum. The momentum of photons, although small, has been measured under laboratory conditions. Use de Broglie's equation to find an equation for the momentum of a photon of wavelength λ.

10 Why can an electron microscope resolve images in finer detail than an optical microscope?

10.3 Light and matter

The idea of wave–particle duality is counter-intuitive and was not immediately accepted by most scientists, even after the ground-breaking work of Einstein, de Broglie and others.

It was the work of Danish physicist Niels Bohr that finally convinced scientists that the particle model was required as part of a complete understanding of the nature of light. Bohr built on the work of Planck and Einstein to explain the emission and absorption spectra of hydrogen (see Figure 10.3.1). This led to important discoveries in astronomy and, eventually, a reformulation of the understanding of the nature of energy and matter.

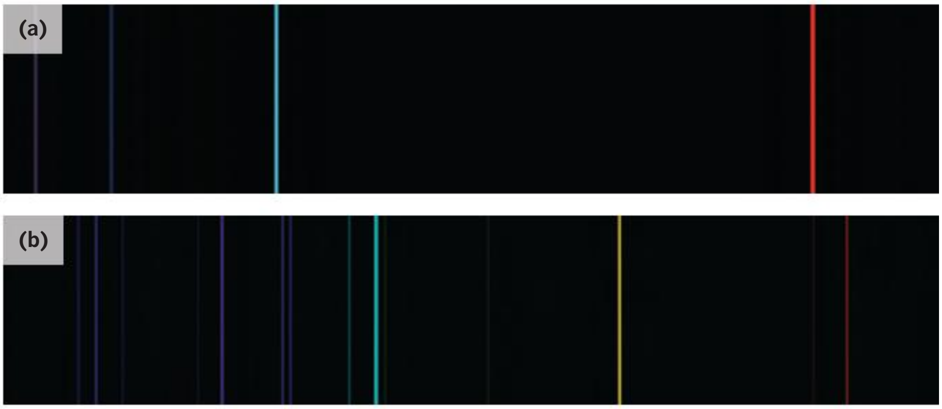


FIGURE 10.3.1 The spectral lines for hydrogen (a) and helium (b). An element's spectral lines are unique to each element and are produced as electrons transition between different energy levels within the atom. Bohr's explanation of emission and absorption spectra was instrumental in furthering the understanding of the nature of energy and matter.

ABSORPTION SPECTRA

In 1814, the German physicist Joseph von Fraunhofer reported a number of dark lines appearing in the spectrum of sunlight, as shown in Figure 10.3.2.

You may recall that a spectrum showing all the colour components of white light can be obtained by passing sunlight through a prism. When Fraunhofer did this, he observed the spectrum (as expected) but also noticed that there were some colours 'missing' from the spectrum. The missing colours appeared as black lines at various points along the spectrum. These apparently missing colours came to be known as Fraunhofer lines.

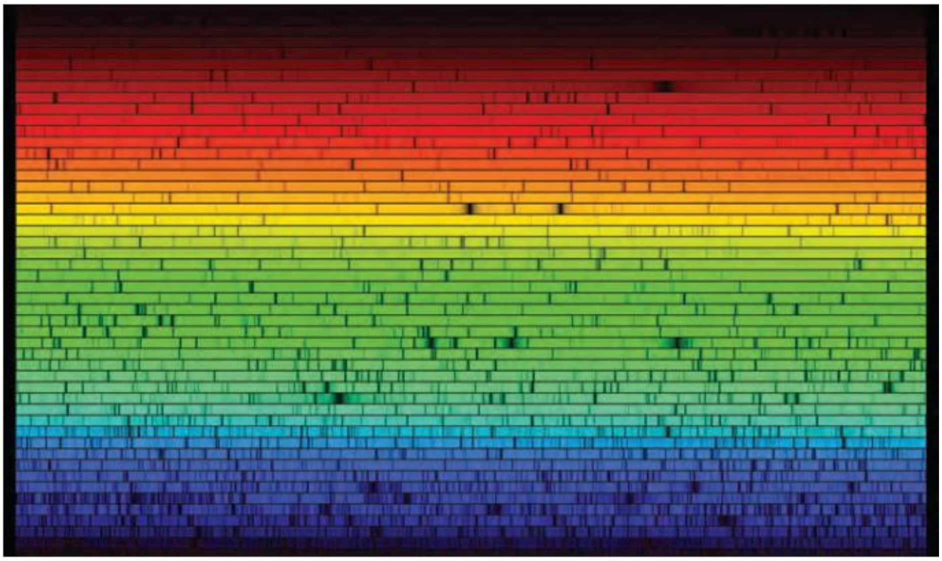


FIGURE 10.3.2 The spectrum of sunlight contains some missing colours known as Fraunhofer lines.

About 50 years later, scientists including Kirchhoff and Bunsen (of Bunsen burner fame) recognised that some of these lines corresponded to the colours emitted when certain gases were heated to high temperatures. They deduced that the dark lines were due to these colours (wavelengths or frequencies) being *absorbed* by gases as light made its way through the outer atmosphere of the Sun. This **absorption spectrum** allowed astronomers to determine that the Sun is largely composed of hydrogen with small quantities of helium and some other heavier elements.

Absorption spectra are valuable for scientists who wish to know what elements are present in a sample of gas or in a solution, so their use is not limited to just astronomy. First, light is directed through a cool sample of a gas or through a solution containing an element or compound. Only certain wavelengths (or frequencies) of light will be absorbed by the elements present in the sample, which means that on viewing the spectrum, this particular wavelength will be 'missing'. The wavelengths that are absorbed are unique to each type of atom. For this reason, by analysing which wavelengths are missing, scientists can determine exactly what elements are present in the sample.

EMISSION SPECTRA

When elements are heated to high temperatures or have an electrical current passed through them, they produce light. Atoms within the material absorb energy and become 'excited' (more on what this means later in this section). This makes the atom unstable and eventually it will return to the 'unexcited' or **ground state**. When this happens, the energy that had been absorbed is released as a single photon. The colour of this photon will depend on the amount of energy it has.

Since atoms can usually have a number of different **excited states**, they can produce a number of different colours. The combination of colours produced by a particular element are distinctive to that element (See Figure 10.3.3) and are known as its **emission spectrum** (shown in Figure 10.3.1 for hydrogen and helium).

FIGURE 10.3.3 The different metals used in fireworks are responsible for the colours in this display. For example, strontium gives red, sodium gives yellow and copper gives green.

SPECTRAL ANALYSIS

In atomic emission spectroscopy, the chemical composition of a material can be determined by analysing the light that is emitted from a material when it is burned or when an electrical current is passed through a gas. The light can be separated into its component wavelengths using a spectroscope, and the specific wavelengths found are characteristic of each particular element, very much like how a fingerprint or DNA is used to identify an individual person.

Recall that an emission spectrum is the result of electrons absorbing energy (the electrons become 'excited') and then releasing energy in the form of a photon. An emission spectrum can be analysed in terms of the energy of the photons produced. In his work on the photoelectric effect, Einstein used Planck's equation for the energy of a photon:

$$\Delta E = hf = \frac{hc}{\lambda}$$

where ΔE is the energy of the photon produced (J)

h is Planck's constant (6.63×10^{-34} J s or 4.14×10^{-15} eV s)

f is the frequency of the photon (Hz)

c is the speed of light (3.00×10^{8} m s^{-1})

λ is the wavelength of the photon (m)

Notice that ΔE has been used in this equation instead of E. ΔE corresponds to the difference in energy between the excited state and the ground state of the electron that released the photon, and so it is used to represent the energy of the photon.

Worked example 10.3.1 relates to the emission spectra of **metal vapour lamps**. Metal vapour lamps produce light as atoms are excited and then emit a photon as they return to their ground state. The emitted photons have wavelengths characteristic of the metals whose atoms are being excited in the lamp. A common type of metal vapour lamp is the sodium lamp. These are often used in street lighting and emit a distinctive yellow colour (see Figure 10.3.4).

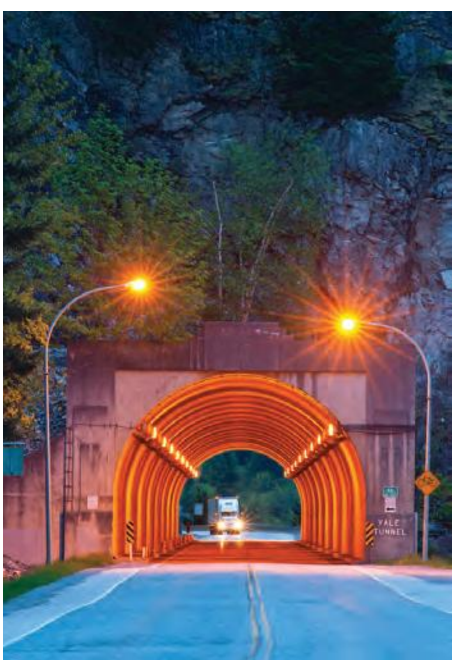


FIGURE 10.3.4 Sodium vapour lamps are commonly used as street lights and have a distinctive yellow colour due to the yellow wavelengths of the sodium emission spectrum.

Worked example 10.3.1

SPECTRAL ANALYSIS

The emission spectrum of a sodium vapour lamp is analysed and shows that most of the light is emitted with a frequency of around 5.1×10^{14} Hz. Calculate the energy of these photons in joules.

Thinking	Working
Recall Planck's equation.	$\Delta E = hf$
Substitute in the appropriate values and solve for ΔE.	$\Delta E = 6.63 \times 10^{-34} \times 5.1 \times 10^{14}$ $= 3.4 \times 10^{-19}$ J

Worked example: Try yourself 10.3.1

SPECTRAL ANALYSIS

In the Sun's absorption spectrum, one of the dark Fraunhofer lines corresponds to a frequency of 6.9×10^{14} Hz. Calculate the energy (in joules) of the photon that corresponds to this line.

The energy of the photon emitted or absorbed can also be expressed in electron-volts (eV) as was explained in Section 10.1. That is, E (in eV) = E (in J) divided by the charge on an electron (1.6×10^{-19} C). To simplify calculations, Plank's constant can be restated in terms of electron-volts, and this figure can be used to calculate the energy directly in electron-volts; i.e. use $h = 4.14 \times 10^{-15}$ eV s.

HYDROGEN'S ABSORPTION SPECTRUM

In the late 19th century, the emission and absorption spectra of hydrogen were of particular interest to scientists as it had been recognised that lines in the absorption spectrum of hydrogen matched lines in the solar spectrum (see Figure 10.3.5).

FIGURE 10.3.5 In the absorption spectrum of hydrogen (a) there is a background of continuous white light (broken into a spectrum of colours), with black lines that correspond to wavelengths of the radiation absorbed by the hydrogen atoms. In the emission spectrum of hydrogen (b) there is a black background against which lines corresponding to the wavelengths emitted by the hydrogen atoms can be observed.

Although some scientists were able to come up with an empirical (based on experimental data) formula that predicted the wavelength of the lines in the hydrogen spectra, no one was able to provide a theoretical explanation for the production of these lines using a wave model for light.

EXTENSION

Balmer and Rydberg—empirical equations

In 1885, the Swiss mathematician Johann Balmer found an empirical equation that predicted the wavelength of the visible lines of the hydrogen emission spectrum:

$$\lambda = \frac{hm^2}{m^2 - n^2}$$

where λ is the wavelength of light (nm)

h is a constant with a value of 365 nm

$n = 2$

m could take values of 3, 4, 5 or 6.

When Balmer put $m = 7$ into the equation, it gave an answer of 397 nm, which corresponded to a spectral line that had been independently observed by Anders Angstrom. Consequently, this set of spectral lines in the visible part of the electromagnetic spectrum came to be known as the Balmer series.

In 1888, Johannes Rydberg (see Figure 10.3.6) realised that Balmer's formula was a special case of the more general formula:

$$\frac{1}{\lambda} = R_H \frac{1}{n^2} - \frac{1}{m^2}$$

where R_H is the Rydberg constant for hydrogen (1.097×10^7 m^{-1})

n and m are any two integers where $m > n$.

This equation predicted that there should be spectral lines in other parts of the electromagnetic spectrum. The ultraviolet series was later observed by Theodore Lyman and two different infrared series were observed by Friedrich Paschen and Frederick Brackett.

FIGURE 10.3.6 Johannes Rydberg developed a general formula predicting the wavelengths of the lines of the emission spectrum of hydrogen.

BOHR MODEL OF THE ATOM

In 1913, the Danish physicist Niels Bohr proposed an explanation for the emission spectrum of hydrogen that drew on the quantum ideas proposed by Planck and Einstein, including Plank's quantum relation equation, $\Delta E = hf$. Bohr realised that:

- The absorption spectrum of hydrogen showed that the hydrogen atom was only capable of absorbing a small number of different frequencies of light and therefore energies of very specific values. That is, the absorbed energy was quantised.
- The emission spectrum of hydrogen showed that hydrogen atoms were also capable of emitting quanta of the exact energy value that it was able to absorb.
- If the frequency, and hence energy, of the incident light was below a certain value the light would pass straight through hydrogen gas without any absorption occurring.
- Hydrogen atoms have an ionisation energy of 13.6 eV. Light of this energy or greater can remove an electron from a hydrogen atom, creating a positive ion.
- Photons of light with all energies above the ionisation value for hydrogen are continuously absorbed.

Bohr's explanation relied on a significant refinement of Rutherford's planetary model of the atom. He devised a sophisticated model of electron energy levels for atoms, a development for which he later won the Nobel Prize in Physics (see Figure 10.3.7).

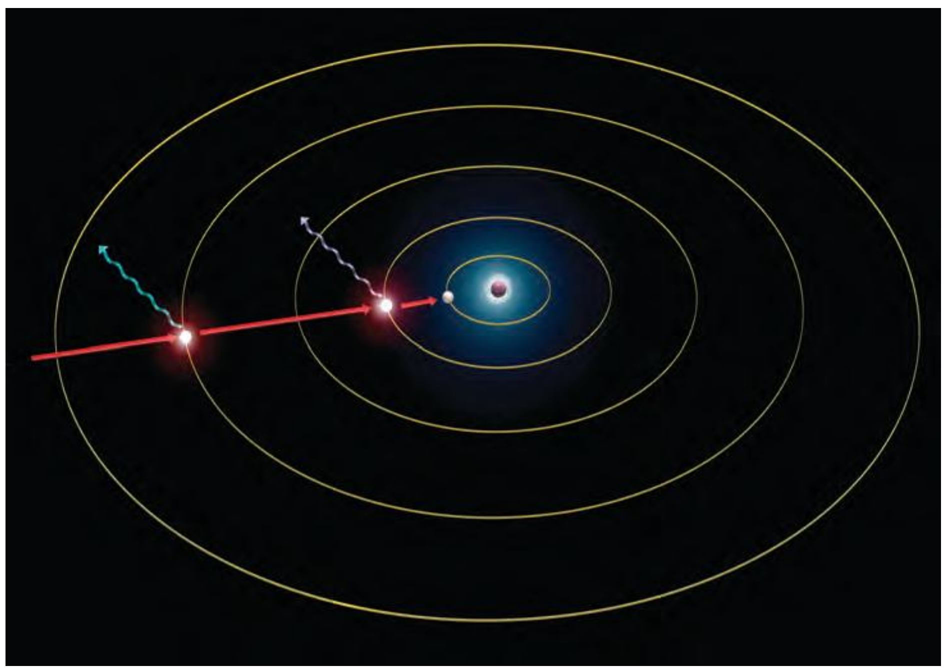


FIGURE 10.3.7 A diagram showing hydrogen spectrum emission levels based on the Bohr model of the atom. Electrons may only orbit in specific energy orbits, shown by the concentric circles. Electrons absorb energy to move to higher levels in their excited states and emit light in specific wavelengths characteristic of the element when returning to the ground state.

Bohr's model of the atom contained the following ideas:

- Electrons move in circular orbits around the nucleus of the hydrogen atom.
- The centripetal force keeping an electron in a circular orbit is the electrostatic force of attraction between the positive nucleus and the negative electron.
- A number of allowable orbits of different radii exist for each atom (labelled $n = 1, 2, 3\ldots$ and known as the principal quantum number). Electrons may only occupy these orbits.
- An electron ordinarily occupies the lowest-energy orbit available (i.e. the ground state).
- An electron does not radiate energy while it is in a stable orbit.

- Electromagnetic radiation (in the form of photons) can be absorbed by an atom when the photon energy is *exactly equal* to the difference in energies between an occupied orbit and a higher orbit.
- Electromagnetic radiation is emitted by an excited atom when an electron returns from a higher energy to a lower energy orbit. The photon energy will be *exactly equal* to the energy difference between the electron's initial and final levels.

Bohr labelled the possible electron orbits for the hydrogen atom with a quantum number (n), and he was able to calculate the energy associated with each quantum number. Using these energy levels, he could theoretically predict the wavelengths of all of the lines of the hydrogen emission spectrum using Planck's equation:

$$\Delta E = \frac{hc}{\lambda}$$

Figure 10.3.8 shows the energy levels for the hydrogen atom. These energies are expressed in terms of how strongly the electron is bound to the nucleus. The ground level ($n = 1$) represents the orbit that is closest to the nucleus, i.e. the unexcited state. An electron in this orbit has an energy of (–)13.6 eV, which means that it would need to gain 13.6 eV of energy for it to escape the atom. Higher energy levels represent orbits that are further from the nucleus.

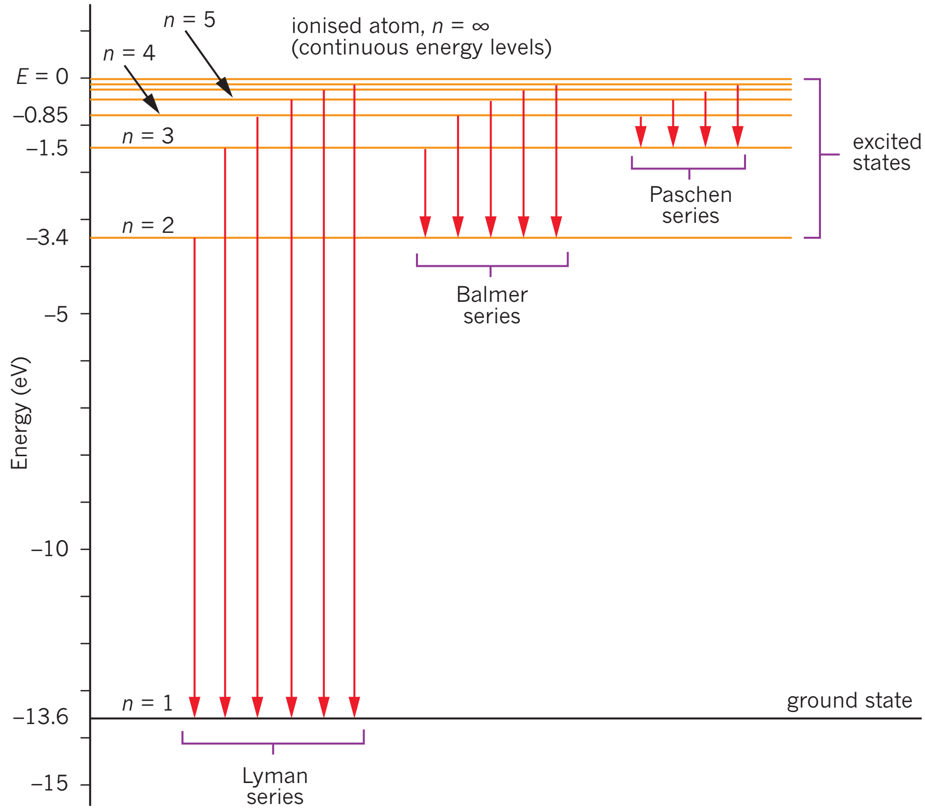


FIGURE 10.3.8 An energy level diagram for hydrogen. An electron in the ground state ($n = 1$) has an energy of (–)13.6 eV. For higher energy levels ($n > 1$), the energy levels can be seen to crowd together.

When a hydrogen atom gains energy, either by heating or from an electrical current, its electron moves from the ground state to one of the higher energy levels. This type of atom is described as 'excited'. Eventually, the electron will drop from the higher energy level to one of the lower levels and will emit a photon with an energy equal to the difference in energy between the levels.

You can see in Figure 10.3.8 that energy levels within the atom are negative in value. A free electron (at $n = \infty$) must have zero potential energy as it has escaped the electrostatic attraction of the proton in the nucleus. To raise an electron from one energy level to another, the appropriate amount of energy must be delivered. As an electron then falls back to its previous energy level, its energy value decreases. That is, it becomes a larger negative number.

Figure 10.3.8 also shows that the spectral lines of hydrogen can be explained in terms of electron transitions. The different series shown on the diagram (Lyman, Balmer, Paschen) represent specific transitions. The Balmer series, for example, shows transitions back to $n = 2$ from various excited energy levels. These transitions represent wavelengths of the visible lines of the hydrogen emission spectrum.

Worked example 10.3.2

USING THE BOHR MODEL OF THE HYDROGEN ATOM

Calculate the wavelength (in nm) of the photon produced when an electron drops from the $n = 4$ energy level of the hydrogen atom to the $n = 2$ energy level. Identify the spectral series to which this line belongs.

Use Figure 10.3.8 to calculate your answer.

Thinking	Working
Identify the energy of the relevant energy levels of the hydrogen atom.	$E_4 = -0.85$ eV $E_2 = -3.4$ eV
Calculate the change in energy.	$\Delta E = E_4 - E_2$ $= -0.85 - (-3.4)$ $= 2.55$ eV
Calculate the wavelength of the photon with this amount of energy.	$E = \frac{hc}{\lambda}$ $\therefore \lambda = \frac{hc}{E}$ $= \frac{4.14 \times 10^{-15} \times 3 \times 10^{8}}{2.55}$ $= 4.87 \times 10^{-7}$ m $= 487$ nm
Identify the spectral series.	The electron drops down to the $n = 2$ energy level. Therefore, the photon must be in the Balmer series.

Worked example: Try yourself 10.3.2

USING THE BOHR MODEL OF THE HYDROGEN ATOM

Calculate the wavelength (in nm) of the photon produced when an electron drops from the $n = 3$ energy level of the hydrogen atom to the $n = 1$ energy level. Identify the spectral series to which this line belongs.

Use Figure 10.3.8 to calculate your answer.

ABSORPTION OF PHOTONS

The Bohr model also explains the *absorption* spectrum of hydrogen (Figure 10.3.5(a) on page 349).

You have already seen that the missing lines in absorption spectra correspond to the energies of light that a given atom is capable of absorbing. This is due to the energy differences between the atom's electron orbits. Only incident light carrying just the right amount of energy to raise an electron to an allowed level can be absorbed.

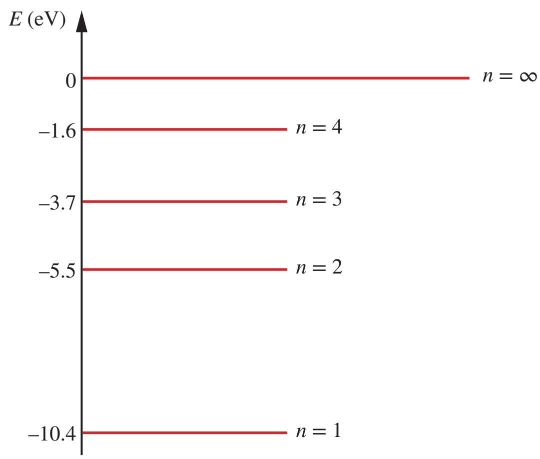
An electron ordinarily occupies the lowest energy orbit. Incident light that does not carry enough energy to raise an electron from this lowest energy level to the next level cannot be absorbed by the atom. Incident light below a certain energy value would simply pass straight through. If light with greater energy than the ionisation energy of an atom is incident, then the excess energy provided by the photon will simply translate to extra kinetic energy for the released electron (recall the photoelectric effect from Section 10.1 of this chapter).

For hydrogen, then, if a hydrogen atom absorbs a photon with 13.6 eV or more, as this is the energy required for the electron to escape the atom completely, the hydrogen atom is said to be 'ionised'.

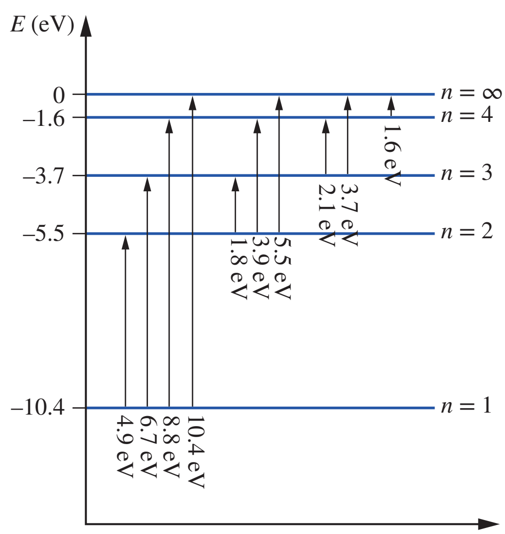
Worked example 10.3.3

ABSORPTION OF PHOTONS

Some of the energy levels for atomic mercury are shown in the diagram below.



Ultraviolet light with photon energies 4.9 eV, 5.0 eV and 10.50 eV is incident on some mercury gas. What could happen as a result of the incident light?

Thinking	Working
Check whether the energy of each photon corresponds to any differences between energy levels by determining the difference in energy between each level.	
Compare the energy of the photons with the energies determined in the previous step. Comment on the possible outcomes.	A photon of 4.9 eV corresponds to the energy required to promote an electron from the ground state to the first excited state ($n = 1$ to $n = 2$). The photon may be absorbed. A photon of 5.0 eV cannot be absorbed since there is no energy level above the ground state that corresponds exactly to 5.0 eV. A photon of 10.5 eV may ionise the mercury atom. The ejected electron will leave the atom with 0.1 eV of kinetic energy.

PHYSICSFILE

The special case of hydrogen

The hydrogen atom was a relatively simple place to begin the development of the field that would come to be known as 'quantum mechanics'. The hydrogen atom contains two charged particles—the positively charged nucleus (which usually contains a single proton) and the electron (see Figure 10.3.9). This means that only one electrical interaction (i.e. between the electron and the nucleus) needs to be considered.

In more complex atoms, such as helium, electrical interactions between the electrons are also significant. This makes the construction of mathematical models for these atoms vastly more complicated than for hydrogen.

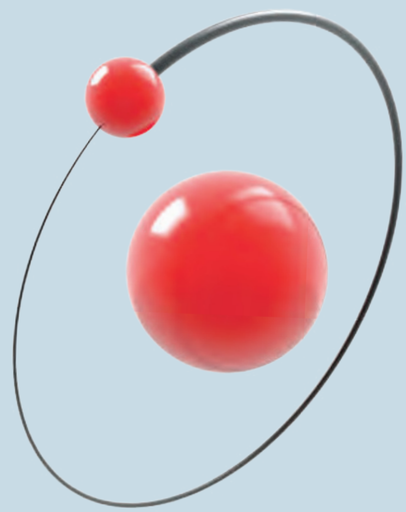
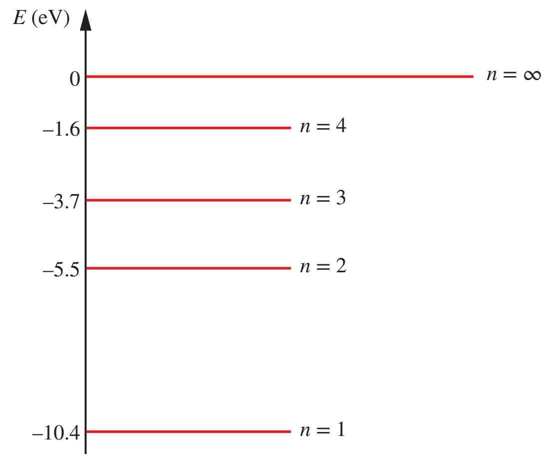


FIGURE 10.3.9 The hydrogen atom contains only two particles: the proton in the nucleus and the electron.

Worked example: Try yourself 10.3.3

ABSORPTION OF PHOTONS

Some of the energy levels for atomic mercury are shown in the diagram below.



Light with photon energies 6.7 eV, 9.0 eV and 11.0 eV is incident on some mercury gas. What could happen to as a result of the incident light?

Problems with Bohr's model

Bohr's model of the hydrogen atom applied a quantum approach to the energy levels of atoms to explain a set of important, previously unexplained phenomena—the emission and absorption spectra of hydrogen. In principle, Bohr's work on the hydrogen atom could be extended to other atoms and, in 1914, the German scientists James Franck and Gustav Hertz demonstrated that mercury atoms contained energy levels similar to hydrogen atoms. Bohr's model signified an important conceptual breakthrough.

However, Bohr's model was limited in its application. It could only really be accurately applied to single-electron atoms—hydrogen and ionised helium. It modelled inner-shell electrons well but could not predict the higher-energy orbits of multi-electron atoms. Nor could it explain the discovery of the continuous spectrum emitted by solids. Further studies even showed problems with the emission spectrum of hydrogen. Some of the observed emission lines could be resolved into two very close spectral lines, and Bohr's model could not explain this. A more complex quantum approach was required.

STANDING WAVES AND THE DUAL NATURE OF MATTER

In the previous section, it was shown that small particles moving at very high speeds can be thought of as matter waves. Wave behaviour can be used to indicate the probability of the path of a particle. If particles can be thought of as matter waves, then these matter waves must be able to maintain steady energy values if the particles are to be considered stable.

De Broglie, the scientist who proposed the idea of matter having wavelengths, applied his approach to the discussion of Bohr's model for the hydrogen atom. He viewed the electrons orbiting the hydrogen nucleus as matter waves. He suggested that the electron could only maintain a steady energy level if it established a **standing wave**.

De Broglie reasoned that if an electron of mass m were moving with speed v in an orbit with radius r, this orbit would be stable if it matched the condition

$$mvr = n\frac{h}{2\pi}$$

where n is an integer.

This can be rearranged to

$$2\pi r = n\frac{h}{mv}$$

Since $2\pi r$ is the circumference, C, of a circle, and using the de Broglie equation, $\lambda = \frac{h}{mv}$, this equation can be rewritten as $C = n\lambda$.

In other words:

The stable orbits of the hydrogen atom are those where the circumference is exactly equal to a whole number of electron wavelengths.

This can be visualised by imagining a conventional standing-wave pattern, like that of a vibrating string discussed in the previous chapter, being looped around on itself in three dimensions as shown in Figure 10.3.10.

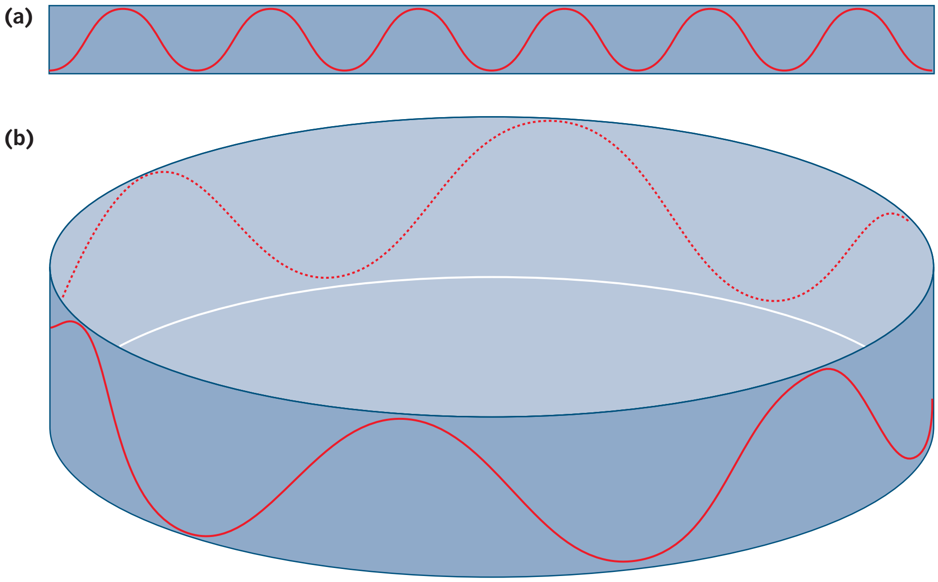


FIGURE 10.3.10 A standing wave pattern (a) can be looped around on itself to form (b) if the circumference of the circle is equal to a whole number of wavelengths.

If the circumference of the circle is not equal to a whole number of wavelengths, then destructive interference occurs, a standing wave pattern cannot be established and the orbit cannot represent an energy level (see Figure 10.3.11).

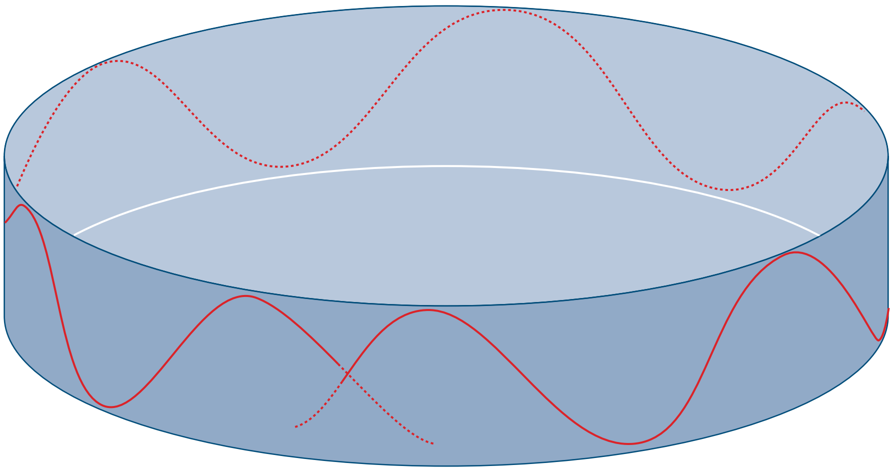


FIGURE 10.3.11 A circular standing wave pattern cannot be established if the circumference of the circle is not equal to a whole number of wavelengths.

COMPARING DIFFERENT LIGHT SOURCES

FIGURE 10.3.12 In an incandescent light globe, electricity is passed through a tungsten filament. As the filament heats up, the free electrons in the tungsten atoms collide, accelerate and emit photons. A wide range of photon wavelengths are emitted due to a wide range of different collisions (some weak, some strong).

So far in this section, you have learnt how photons of light are emitted when electrons move from an excited state to ground state. But this is not the only way to produce light. An **incandescent** light bulb produces light by heating a filament to a very high temperature (see Figure 10.3.12). This produces electromagnetic radiation at a range of wavelengths; that is, the light from the incandescent globe is a continuous spectrum. Some of the light produced is actually in the infrared part of the spectrum that is invisible to human beings (see Figure 10.3.13).

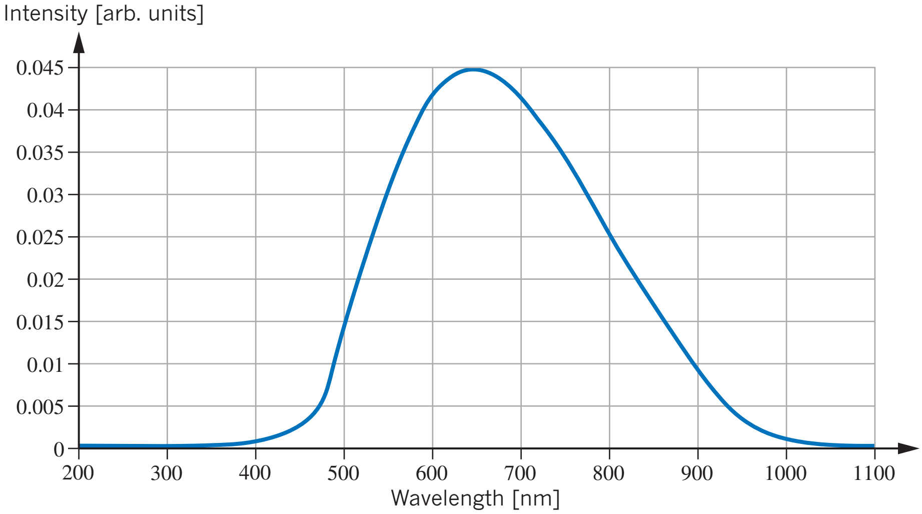


FIGURE 10.3.13 The spectrum of a 230 V, 60 W incandescent light globe. Since the visible spectrum is between about 400 nm and 750 nm, some of the radiation produced is not visible to humans.

PHYSICSFILE

Inefficiency

Incandescent light bulbs are widely seen as one of the most inefficient forms of lighting available. Traditional incandescent bulbs convert less than 10% of the electrical energy into light, with the remainder being released as heat and other forms of long wavelength electromagnetic radiation. As a consequence, these types of light bulbs are being phased out and replaced with compact fluorescent lights (CFLs), which are much more efficient and longer lasting.

Other light sources emit light with different properties due to the different methods employed to produce them.

Coloured LEDs

FIGURE 10.3.14 LEDs can be produced in virtually every visible and near-visible colour.

A **light-emitting diode** (LED) is a semi-conducting device that uses the excitation of electrons to produce light (see Figure 10.3.14). Most LEDs are made primarily from silicon. Pure silicon is a relatively poor conductor. However, the addition of a small amount of another material, in a process known as 'doping', can change the conductive properties of silicon dramatically.

A semi-conductor diode is designed so that most electrons sit in an energy level known as the *valence band*. At a slightly higher energy is another energy level known as the *conduction band*. If a small amount of electrical energy is provided as a potential difference, electrons can jump from the valence band to the conduction band and then move through the silicon as an electric current. When the electrons eventually drop back into the valence band, their energy is released as photons.

The colour of the light produced is determined by the energy gap between the valence and conduction bands. This in turn depends on the amount and type of doping of the silicon. This means that scientists can effectively 'tune' LEDs to produce photons of a particular wavelength, and hence to produce light of a particular colour.

Lasers

To produce light from gas, the atoms in the gas need to be raised to an 'excited state' by heating the gas or passing an electric current through it. In a normal sample of gas, most of the gas atoms are in the ground (unexcited) state and the relatively small number of excited atoms emit their photons of energy spontaneously and in no fixed pattern.

However, it is possible to create a system known as a 'population inversion', in which most of the gas atoms are in an excited state. If a photon of the appropriate energy is then introduced to this gas, it can stimulate these atoms to release their photons in a systematic way. This is the basis of the **laser** (see Figure 10.3.15).

The term laser is an acronym that stands for *Light Amplification by Stimulated Emission of Radiation* (LASER). Laser light has a number of unique properties. It is usually polarised and either monochromatic or limited to a very narrow band of wavelengths. Laser light is also coherent, which means that all of the waves are in phase, i.e. their crests and troughs occur in time with each other.

Although lasers were first developed as a product of theoretical research, they now have a wide variety of applications in areas as diverse as communications, medicine and weaponry.

FIGURE 10.3.15 Laser light is coherent, polarised and monochromatic.

Synchrotron light

A **synchrotron** is a machine that uses powerful magnets to accelerate charged particles, usually electrons, to velocities close to the speed of light (see Figure 10.3.16). You may have already come across this idea in the optional material in Unit 2 or be aware of the Australian Synchrotron based near Monash University in Melbourne.

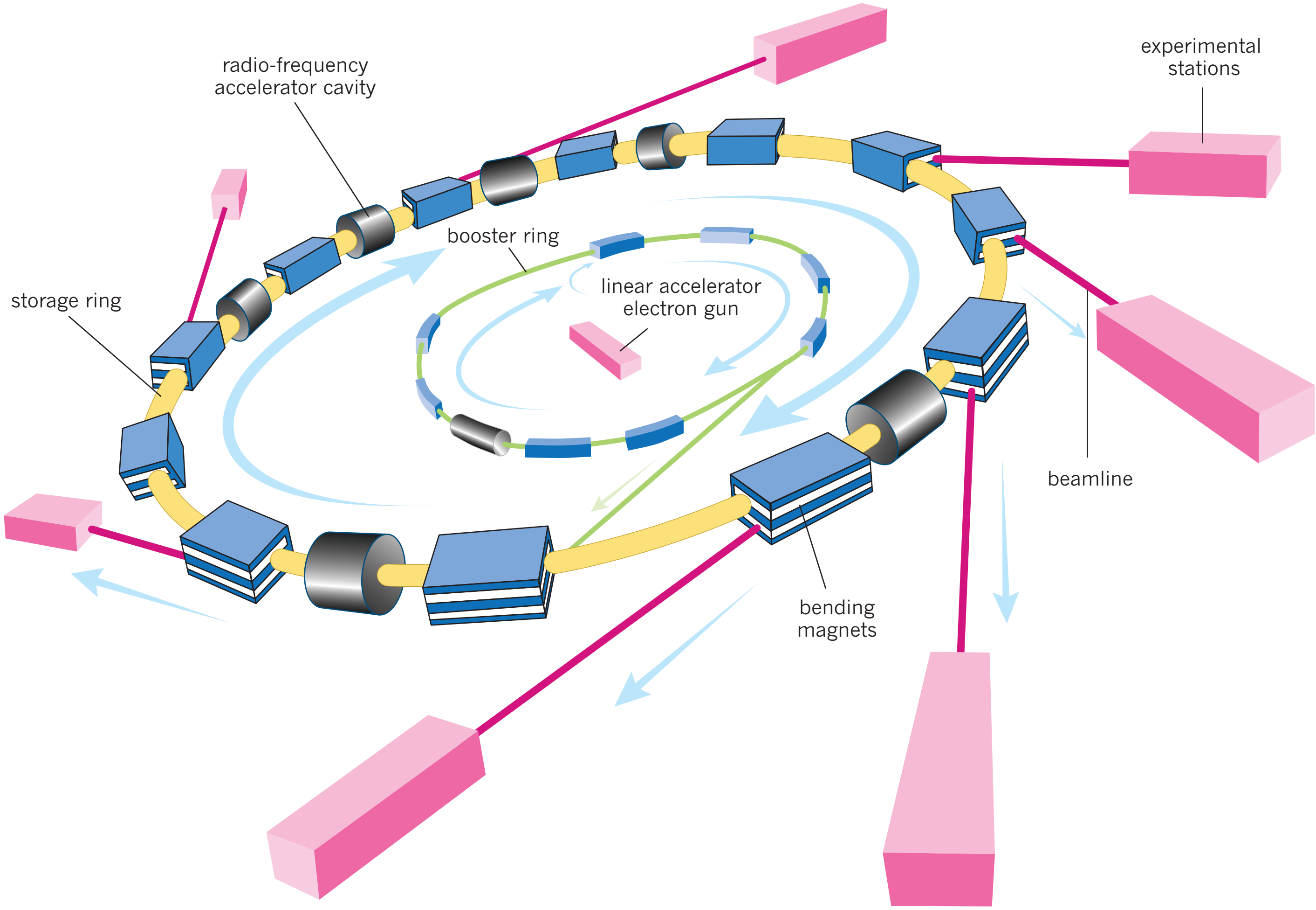


FIGURE 10.3.16 A synchrotron accelerates electrons to near light speed.

Most synchrotrons are very large so that the electrons can be accelerated to the very high speeds required. For example, the Australian Synchrotron in Clayton, Victoria has a circumference of over 200 m.

When electrons are accelerated, they produce electromagnetic radiation (light) known as synchrotron light. Because of the extremely high energies involved, the electromagnetic radiation produced by a synchrotron has a number of special properties. It is:

- extremely bright
- highly polarised
- emitted in very short pulses
- produced across a broad range of wavelengths from microwaves to gamma rays.

Synchrotron light has a wide range of scientific uses across medicine, bioscience, materials science and engineering. For example, synchrotron light is useful for exploring the structure of very small objects, smaller than those that can be seen with visible light. An ordinary light microscope is incapable of resolving many small structures due to the long wavelength of visible light. On the other hand, the short-wavelength X-rays in synchrotron light are ideal for examining structures at the cellular or atomic level as they can resolve images down to the size of individual atoms (see Figures 10.3.17).

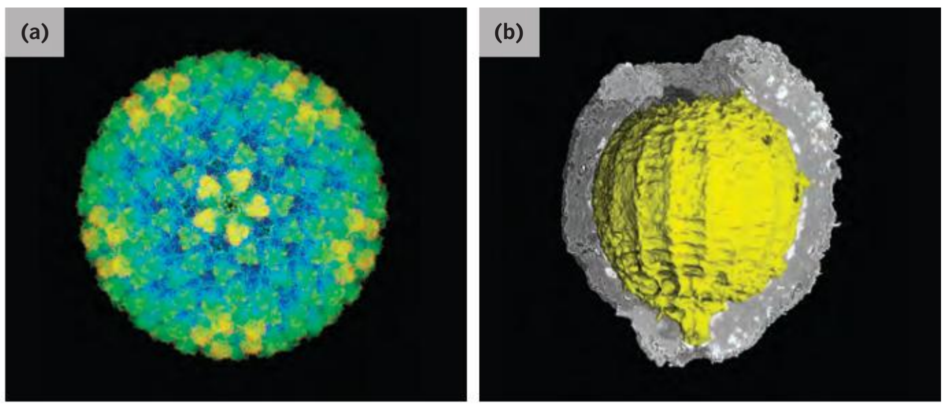


FIGURE 10.3.17 (a) Colour-enhanced image of the Bluetongue virus obtained from the diffraction pattern of high-energy X-rays from synchrotron radiation. The virus gets its name for the blue tongue it causes in sheep. (b) Synchrotron light was used to map the internal structure of this fossilised alga without having to cut or break it open.

10.3 Review

SUMMARY

- The production of spectra suggests an internal structure to the atom. A line *emission* spectrum is produced by energised atoms, an *absorption* spectrum is created when white light passes through a cold gas.
- The spectrum for any particular element is unique to that element.
- Bohr suggested that electrons in atoms orbit the nucleus in specially defined energy levels. No radiation is emitted or absorbed unless the electron can jump from one energy level to another. Electron energies are said to be *quantised*, since only certain values are allowed.
- The frequency of a photon emitted or absorbed by a hydrogen atom can be calculated from the difference between the energy levels involved, i.e. $E_2 - E_1 = hf$ or $= \frac{hc}{\lambda}$.
- The Bohr model of the atom is limited in its application, but was a significant development at the time as it took a quantum approach to the energy levels of atoms and incorporated the quantum nature of electromagnetic radiation.
- de Broglie viewed electrons as matter waves. His standing-wave model for electron orbits provided a physical explanation for electrons only being able to occupy particular energy levels in atoms. He suggested that the only way that the electron could maintain a steady energy level was if it established a standing wave.
- The quantised states of the atom are analogous to the quantised standing waves that are known to occur in physical objects such as strings.
- In an incandescent lamp, the thermal motion of free electrons produces a continuous spectrum. Lasers, LEDs, metal vapour lamps and synchrotrons are light sources that produce light at discrete frequencies via the emission of photons when excited electrons release energy. The means of excitation varies by source.

KEY QUESTIONS

1 When does an element such as sodium produce an emission spectrum?

2 An emission line of frequency 6.0×10^{14} Hz is observed when looking at the emission spectrum of a particular elemental gas. What is the energy, in joules, of photons corresponding to this frequency?

3 Photons of energy 0.42 eV are emitted by a particular atom as it returns from the excited to the ground state. What is the corresponding wavelength of these photons?

4 What do the following acronyms stand for?
  a LED
  b LASER

5 At what wavelength (in nm) will an LED radiate if it is made from a material with an energy band gap of 1.84 eV?

6 Calculate the energy of the photon required to move an electron in a hydrogen atom from its ground state (n = 1) to the n = 4 energy level. Refer to Figure 10.3.8.

7 Calculate the wavelength of the photons released in the transition described in Question 6.

8 What do de Broglie's matter-wave concept and a bowed violin string have in common?

9 Bohr's quantised model of the atom was a significant development. However, it was limited in application. What was the Bohr model unable to explain?

10 When an electron drops from the n = 5 energy level of the hydrogen atom to the n = 2 energy level, a 434 nm photon is released. If the n = 2 orbit has an energy of –3.4 eV, what is the energy of the n = 5 orbit?

10.4 Heisenberg's uncertainty principle

In the early 20th century, scientists struggled to interpret the evidence of the dual wave–particle nature of energy and matter. Waves and particles are fundamentally different—waves are extended and continuous whereas particles are discrete. How two such different models could be combined to describe the fundamental building blocks of nature was a serious puzzle. As scientists delved into this mystery, they discovered fundamental limitations to their ability to explore the 'quantum' universe.

A QUANTUM INTERPRETATION OF THE ELECTRON

In 1925, the Austrian physicist Erwin Schrödinger built on the work of Niels Bohr by developing a mathematical equation that could describe the wave behaviour of electrons in situations other than the simple hydrogen atom.

In Schrödinger's model, the wave properties of electrons are interpreted as representing the probability of finding an electron in a certain location.

$$i\hbar \frac{\partial}{\partial t}\Psi = \hat{H}\Psi$$

FIGURE 10.4.1 Schrödinger's wave equation. While outside the scope of this course, it is interesting to see what Schrödinger's equation looks like. The Greek symbol Ψ (psi) represents the wave function of the electron.

Quantum mechanics is the name now given to the area of physics in which the wave properties of electrons are studied. Schrödinger's equation (see Figure 10.4.1) has been used to calculate the regions of space in which an electron can be found in a hydrogen atom. These are now known as 'orbitals' rather than orbits because they are complex three-dimensional shapes, as shown in Figure 10.4.2, rather than the simple circular paths once imagined by Rutherford and Bohr.

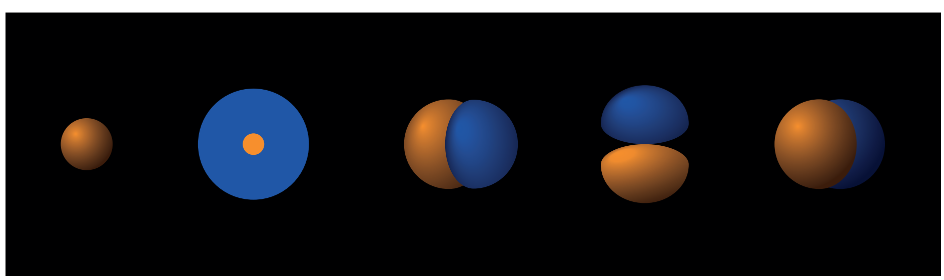


FIGURE 10.4.2 The shapes of the first five electron orbitals of a hydrogen atom.

PHYSICSFILE

Schrödinger's cat

Schrödinger described the strange, counter-intuitive nature of quantum mechanical systems using a now-famous analogy known as Schrödinger's cat.

This is a thought experiment (i.e. Schrödinger did not actually perform the experiment) in which a cat is placed in a closed box with a flask of poison. A quantum mechanical system is set up such that there is a 50% chance of the flask being broken and the cat killed.

Schrödinger argued that until the box is opened to reveal the outcome of the experiment, the cat is considered as simultaneously alive *and* dead.

In a manner similar to the dual nature of light, the outcome (for the cat being alive or dead; for light being a wave or a particle) does not exist until an observation or measurement is made.

LIMITS TO MODELS AT VERY SMALL SCALES

It was becoming clear to scientists that the nature of the universe at the very smallest scale is fundamentally different to the way the universe is perceived at the macroscopic scale.

In everyday life, each object has a clearly definable position and motion. The classical laws of physics, developed by scientists from Newton through to Maxwell, are all based on this assumption, which is so fundamental to human experience that it is hard to imagine a universe where this is not the case.

However, this is exactly what is needed in order to explore the quantum universe. There is no particular reason why tiny particles such as electrons and photons should be similar to larger objects like balls or planets; scientists initially just extrapolated from their experience until the evidence showed that their assumptions were wrong.

HEISENBERG'S UNCERTAINTY PRINCIPLE

Whenever a measurement is taken a degree of error or *uncertainty* is involved. The certainty of the measurement is limited by the measuring resolution of the device used to make the measurement. For example, a ruler with markings one millimetre apart will have an uncertainty in any measurement of about half a millimetre. This is taken into account when commenting on the errors in a practical experiment or calculating the final uncertainty in a result. If more precision is needed in a final result, then a more precisely marked measuring device such as a micrometer or vernier calliper is needed to make the initial measurements. The smaller the divisions, the smaller the final uncertainty in the result.

However, according to quantum mechanics there is a physical limit to the absolute accuracy of particular measurements. This limit is inherent in nature and is a result of both wave–particle duality and the interactions between the object being observed and the effect of the observation on that object (as Schrödinger tried to explain). The first scientist to clearly identify this limit was the German physicist Werner Heisenberg (see Figure 10.4.3). The **Heisenberg uncertainty principle** describes a limit to which some quantities can be measured.

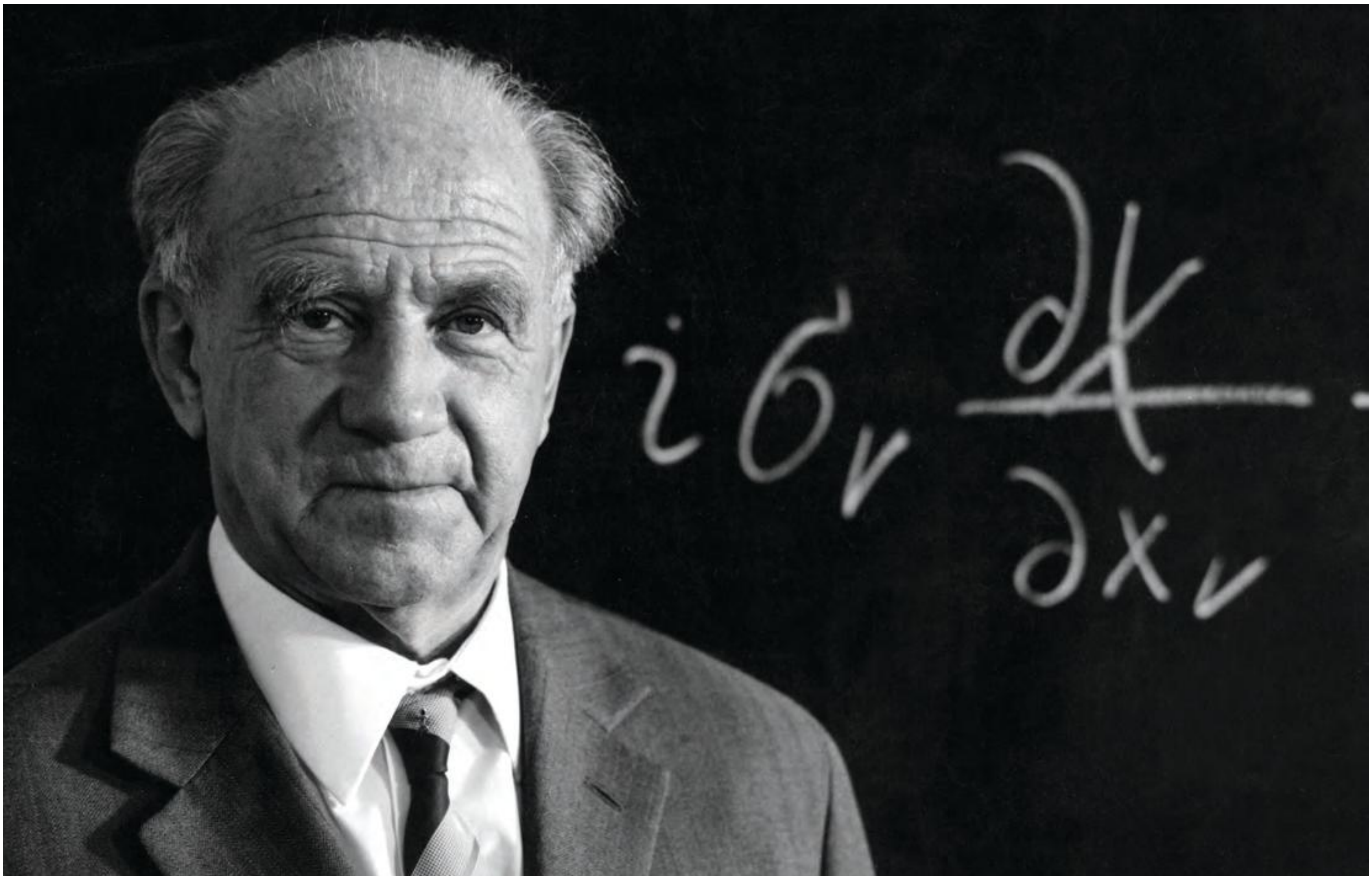


FIGURE 10.4.3 Werner Heisenberg won the Nobel Prize in 1932 for his work on the uncertainty principle. Heisenberg is regarded as a founder of quantum physics.

Imagine trying to find a ball in a pitch-black room. The only way to do so, assuming there is no light, is to feel around. The more you search and feel around, the more confident you can be that the object remains in the area of the room yet to be searched—until you actually touch it. Then there is every chance that it will roll away, the ball having been given momentum by your touch. You'll no longer know its *future* position. The very act of determining the position of the ball made knowing its future position less certain.

Similarly, at the quantum level, to measure the exact location of an electron it would be necessary to hit the electron with another particle such as a photon of light. However, as soon as the photon strikes the electron, it would cause the electron to move as the photon transfers some or all of its energy to the electron. The act of measurement causes a change in the value being measured. This is a general problem when trying to measure the location or motion of all sub-atomic particles.

According to Heisenberg's uncertainty principle, the more exactly the position of a sub-atomic particle is known, the less is known about its momentum. Similarly, the more precisely the momentum of a particle is measured, the less certain is its exact position.

Heisenberg went on to explain his uncertainty principle with a formula that states that the product of the uncertainties in the position (Δx) and momentum (Δp) of a particle must always be greater than a certain minimum value related to Planck's constant. That is:

$$\Delta x \Delta p \geq \frac{h}{4\pi}$$

or alternatively in terms of energy (since the photon may transfer some or all of its energy to the electron):

$$\Delta E \Delta t \geq \frac{h}{4\pi}$$

where ΔE is the uncertainty in the energy of an object

Δt is the uncertainty in the time taken for the energy transfer to occur.

You may see this equation stated differently in some references. The difference is due to the way in which the variables are defined. The important part of the relationship is the fact that Δx and Δp are inversely proportional. When either Δx or Δp increase, the other must decrease.

This equation is sometimes called the 'indeterminacy principle'. It states that it is not possible to know both the position and the momentum of an object at exactly the same time (see Table 10.4.1). The more accurately the position is measured, the greater the uncertainly in the momentum and vice-versa. Note that this doesn't infer that an absolute measure of position cannot be made. Just that, in doing so, its momentum at that same time wouldn't be known and hence there is no way of knowing what the position would be a second later.

	Position	Momentum
Scenario 1	known	unknown
Scenario 2	unknown	known

TABLE 10.4.1 It is not possible to accurately know both the position and the momentum of an object at exactly the same time.

While, in the examples above, the position of an object or actual particle has been considered, Heisenberg's uncertainty principle applies particularly at the quantum level precisely because electrons aren't particles. The whole dual nature of matter idea means that electrons, and in fact *all* matter, behave *both* as a wave and as a particle. The uncertainty principle applies a limit to the simplified idea of electrons as particles. i.e. that the position and the velocity of an electron cannot both be known at exactly the same time, and that the amount of energy at a particular time, t, is also uncertain. For the normal-sized world around us, the inclusion of Planck's constant, h, in the measure of uncertainty means that the level of uncertainty in determining the position of everyday objects is extremely small—in fact, virtually insignificant.

However, at an atomic scale, this level of uncertainty is substantial. And since everyday objects are made up of atoms containing sub-atomic particles such as electrons, the basic understanding of matter comes down to this fundamental property of all quantum mechanical systems.

EXTENSION

Examples

It is not necessary in this course to be able to calculate the uncertainty in the position of an electron. However, some simple examples illustrate the difference in proportions of uncertainty at the quantum and everyday world levels.

Assuming an electron is moving with a speed of 1.0×10^6 m s^{-1}, its momentum would be:

$p = mv = 9.11 \times 10^{-31} \times 1.0 \times 10^6 = 9.11 \times 10^{-25}$ kg m s^{-1}

If the momentum is known with an uncertainty of 1% then $\Delta p = 9.11 \times 10^{-27}$ kg m s^{-1}.

Using the uncertainty principle, the uncertainty in position would be:

$$\Delta x \Delta p \geq \frac{h}{4\pi}$$

$$\Delta x \geq \frac{h}{4\pi \Delta p} = \frac{6.63 \times 10^{-34}}{4\pi \times 9.11 \times 10^{-27}} = 5.8 \times 10^{-9} = 5.8 \text{ nm}$$

Since the diameter of an atom is between 0.1 nm and 0.5 nm, the uncertainty in knowing the position of the electron is many times the diameter of an atom!

Applying the same process to determining the uncertainty in position of a 600 g basketball travelling at 10 m s^{-1}:

$p = mv = 0.6 \times 10 = 6$ kg m s^{-1}

If the momentum is known with an uncertainty of 1% then $\Delta p = 0.06$ kg m s^{-1}.

Using the uncertainty principle, the uncertainty in position would be:

$$\Delta x \Delta p \geq \frac{h}{4\pi}$$

$$\Delta x \geq \frac{h}{4\pi \Delta p} = \frac{6.63 \times 10^{-34}}{4\pi \times 0.06} = 8.8 \times 10^{-34}$$

This is a barely measureable uncertainty and means that the position of the basketball is very easily predicted. The uncertainty principle sets no real limit to the measurement of everyday objects.

PHYSICS IN ACTION

Viewing an electron

Imagine trying to view an electron with an optical microscope like that shown in Figure 10.4.4 (the situation would be similar with an electron microscope). For the electron to be seen, a photon would have to strike the electron and be reflected back to the observer. As noted earlier in this chapter when looking at X-ray diffraction, objects can be seen at their best when the wavelength of the electromagnetic radiation used is at least as small as the object. Since a short wavelength corresponds to a high frequency and high energy (i.e. $E = hf = \frac{hc}{\lambda}$), the photons needed to observe the electron would have high energy and thus would impart more momentum to the object being observed. The higher the energy, the shorter the wavelength and the better the potential resolution, but the more likely it would be that the photon would knock the electron off course and hence the object's position would be subject to greater uncertainty. Just attempting to observe the electron introduces significant uncertainty in either the position or the momentum of the electron.

How then can an electron be viewed?

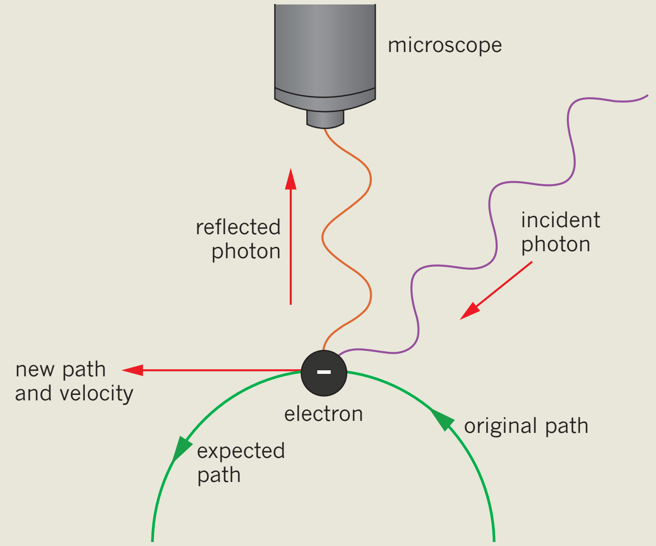


FIGURE 10.4.4 A thought experiment considering how an electron could be observed. The reflection of the photon needed to observe the electron introduces uncertainty in the position of the electron, making it unobservable.

PHYSICSFILE

Interpretation

Scientists have long argued about how we should interpret Heisenberg's uncertainty principle. In the early 20th century, many were unhappy with the implication of Schrödinger's work that the motion of an electron could only be described in terms of probabilities. Even Einstein famously argued against this interpretation, saying 'God does not play dice with the universe' (see Figure 10.2.9).

However, experiments over the last century have confirmed that uncertainty is a fundamental property of the quantum mechanical universe.

FIGURE 10.4.5 Einstein's claim that 'God does not play dice' is challenged by modern understandings of quantum physics.

SINGLE-SLIT DIFFRACTION AND THE UNCERTAINTY PRINCIPLE

An experiment that can be used to illustrate Heisenberg's uncertainty principle is the single-slit diffraction of light.

If light from a laser is shone through a narrow adjustable slit, a diffraction pattern will form on the screen behind the slit (see Figure 10.4.6). This idea was covered in an earlier section.

This diffraction pattern can be explained by treating the laser light as a wave and considering the interference of different sections of the wavefront that passes through the slit. However, explaining this pattern in terms of the motion of individual photons is much more challenging.

The experiment has been conducted using a light so dim that it can be reasonably inferred that one photon passes through the slit at a time. This rules out the possibility that the diffraction pattern could be caused by photons interacting with each other.

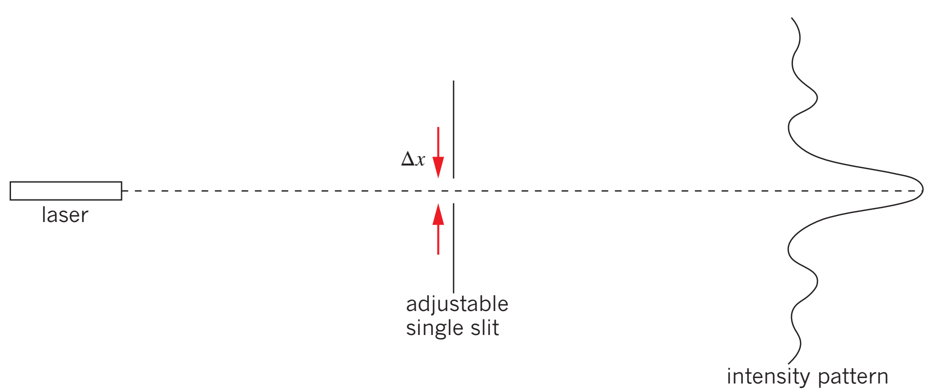


FIGURE 10.4.6 Demonstrating Heisenberg's uncertainty principle with diffraction through a single slit.

Consider now a single photon on its journey from the laser to the screen. As the photon passes through the slit, its position is known to some degree of certainty (Δx). According to Heisenberg, this means that there must therefore be some uncertainty about the photon's momentum (Δp), which is why the photon can end up at a variety of different places on the screen. Schrödinger's wave equation could be used to explain why some paths for the photon are much more likely than others.

Since the slit is adjustable, it could be made narrower. This would decrease the uncertainty about the photon's position (i.e. Δx would decrease) so the uncertainty about the photon's momentum would increase (i.e. Δp would increase). This is why narrowing the slit causes the diffraction pattern to spread out.

Compare this with the same experiment with the slit removed, as shown in Figure 10.4.7.

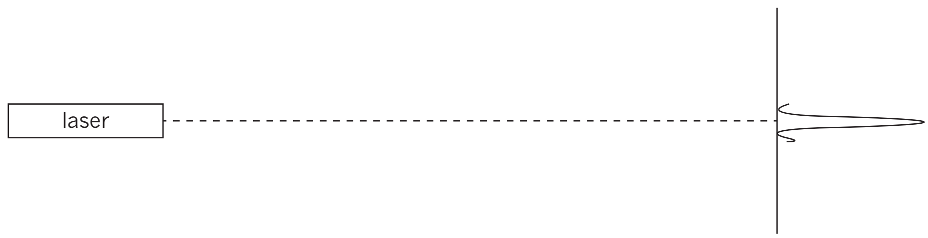


FIGURE 10.4.7 Without a slit, there is no uncertainty about the momentum of the photon and therefore no diffraction occurs.

Since there are no constraints on the path of the light, the uncertainty about the position of each individual photon is large. Correspondingly, the uncertainty about the momentum of the photon becomes is very small and no diffraction occurs—all the photons end up very close together in a small spot in the middle of the screen.

A QUANTUM VIEW OF THE WORLD

Bohr's model of the atom, explained in Section 10.3, was a mixture of classical and quantum theories, partially recognising the wave–particle duality of light and matter. However, it allowed for the development of new, more-radical theories to be developed by physicists such as Schrödinger and Heisenberg. And intriguingly, quantum mechanics has confirmed certain aspects of the Bohr model, such as atoms existing only in discrete states of definite energy and the emission or absorption of photons of light when electrons make transitions from one energy state to another within an atom.

However, quantum mechanics goes much further—very much further than the brief introduction in this section. According to quantum mechanics, electrons don't exist in well-defined circular paths as we so regularly depict them in books. Because electrons are not particles, they don't follow particular paths in space and time at all. Rather, because of the wave nature of an electron, the paths can be thought of more as clouds, where the particular location of an electron at any point in time is based on probability since to measure a precise point, according to Heisenberg, only introduces uncertainty preventing you from knowing where the electron would be at the next moment in time (see Figure 10.4.8).

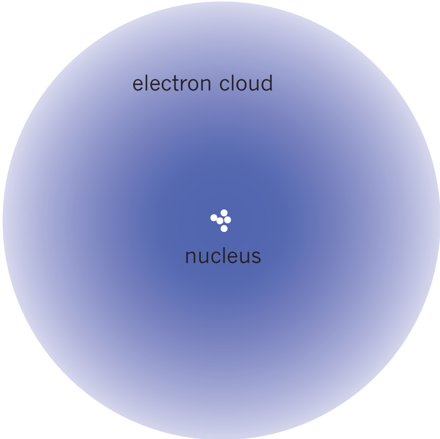


FIGURE 10.4.8 An atom as quantum mechanics views it. This conceptual image depicts the position of an electron around an atomic nucleus as an electron cloud or probability distribution of the location of the electron at any point in time.

The classical view of the world is a Newtonian one, where once the position and speed of an object is known, its future position can be predicted. This is termed a 'deterministic' model and works particularly well in predicting the positions of ordinary objects. This model suggests that the future of the universe, made of up of particle-like objects, is completely knowable. Quantum mechanics proposes something very different where the dual nature of particles, particularly fundamental particles such as the electron, prevents knowing the position and speed of an object at the same time. We can only calculate the probability that an electron will be observed at a particular place around an atom. In this view of the world there is some inherent unpredictability. In fact, it becomes meaningless to ask *how* an electron gets from one state to another when an atom emits or absorbs a photon of light—it just does.

10.4 Review

SUMMARY

- The nature of the universe at the very smallest of scales is fundamentally different to the way the universe is perceived at the macroscopic scale.
- Quantum mechanics is the study of the wave properties of electrons. These wave properties are interpreted as describing the probability of finding an electron at a particular point in space.
- Heisenberg's uncertainty principle results from wave–particle duality and states that it is not possible to know the exact position and momentum of a particle simultaneously. The more exactly the position of a sub-atomic particle is known, the less is known about its momentum. Similarly, the more precisely the momentum of a particle is measured, the less certain is its exact position.
- The single-slit diffraction experiment provides an example of the application of the uncertainty principle. The diffraction pattern is not produced by photon interactions but rather the probability of where a single photon may end up. Some positions are more likely than others, hence there is a larger central maximum intensity.
- In the quantum mechanical view of the atom, electrons are not particles and do not have clearly defined orbits. Their paths, due to their wave nature, can be explained as a probability distribution of their positions as particles or an electron cloud.

KEY QUESTIONS

1 In a particular experiment, if the uncertainty about the position of a particle were decreased, then what will happen to the uncertainty about the speed of the particle?

2 Heisenberg's uncertainty principle was a contributing factor in showing Bohr's model of the atom to be inaccurate. How does the uncertainty principle contribute to the inaccuracy of Bohr's model?

3 Explain why Newtonian (classical) laws of physics are not appropriate when describing what occurs at the sub-atomic level.

4 Consider the situation of using an optical telescope to view an electron. In order to see the electron, a photon is reflected from it back to the observer. Describe what would happen when the photon collides with a moving electron.

5 The position of an everyday object can be readily predicted yet that of an electron cannot. Why is this?

6 Single-slit diffraction provides evidence of the dual nature of light and so can be explained by the uncertainty principle. What happens when the slit is removed?

7 The single-slit diffraction experiment produces the same results if it is performed using electrons instead of photons. What happens when the slit width is increased in this experiment?

8 When is it appropriate to apply the Heisenberg uncertainty principle?

9 Which of the following statements is *not* consistent with Heisenberg's uncertainty principle?

A The position and momentum of a particle can never be exactly known at the same time.

B The more exactly the position of particle is measured, the more uncertainty there is about its momentum.

C If the momentum of a particle is precisely measured, its position will be unknown.

D It is possible to precisely measure the momentum and position of a particle simultaneously.

Chapter review

10

KEY TERMS

absorption spectrum
de Broglie wavelength
electron-volt
emission spectrum
excited state
ground state
Heisenberg's uncertainty principle
incandescent
laser
light-emitting diode
metal vapour lamp
photocurrent
photoelectric effect
photoelectron
photon
quantum
quantum mechanics
standing wave
stopping voltage
synchrotron
threshold frequency
wave–particle duality
work function

1 What is the energy, in electron-volts, of light with a frequency of 6.0×10^{14} Hz?

2 What is the approximate value of the energy in J of a quantum of light with energy of 5.0 eV?

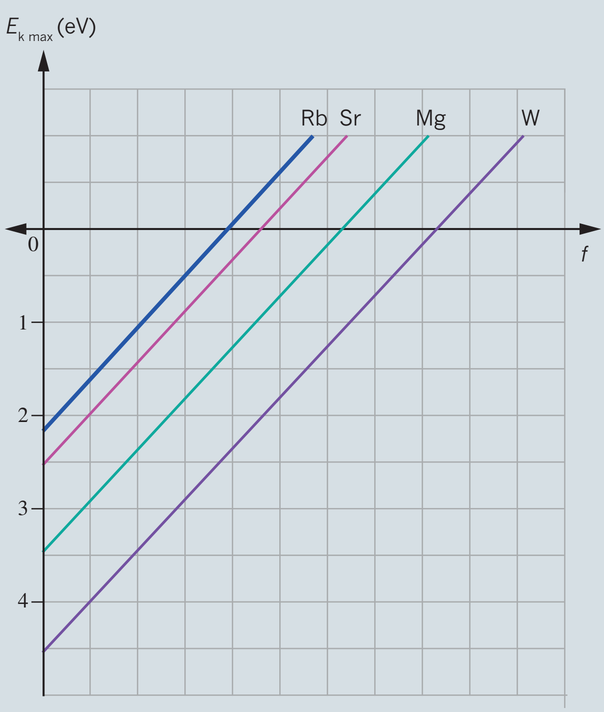
3 What name is given to the electrons released from a metal surface due to the photoelectric effect?

4 If the work function for nickel is 5.0 eV, what is the threshold frequency for nickel?

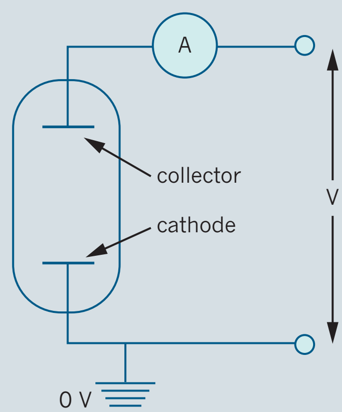
5 Platinum has a threshold frequency of 1.5×10^{15} Hz. Calculate the maximum kinetic energy, in electron-volts, of the emitted photoelectrons when ultraviolet light with a frequency of 2.2×10^{15} Hz shines on it.

6 The stopping voltage obtained using a particular photocell is 1.95 V. Determine the maximum kinetic energy of the photoelectrons in electron-volts.

7 From the graph, determine the value of the work function for each of the metals.



8 The cathode of a particular photocell, shown below, is coated with rubidium. Incident light of varying frequencies is directed onto the cathode of the cell and the maximum kinetic energy of the photoelectrons is logged. The results are summarised in the following table.



Frequency (Hz) × 10^{14}	$E_{k\ max}$ (eV)
5.20	0.080
5.40	0.163
5.60	0.246
5.80	0.328
6.00	0.411
6.20	0.494

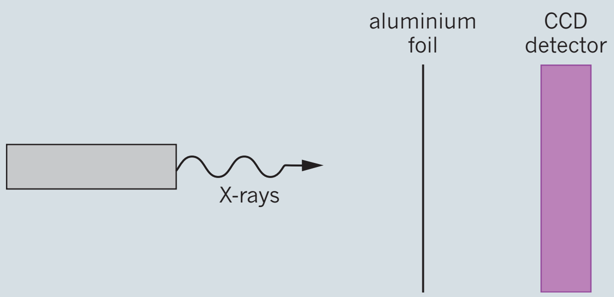
a Plot the points from the table on a graph.

b Calculate the gradient of the graph.

c Based on the graph of the experimental results, what is the threshold frequency for rubidium?

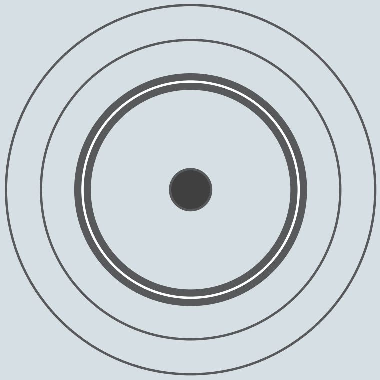
d Will red light of wavelength 680 nm cause photoelectrons to be emitted from the rubidium surface? Justify your answer.

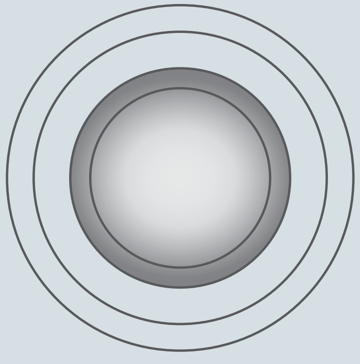
9 In an X-ray diffraction experiment, a beam of X-rays from a synchrotron is directed on to a sheet of thin aluminium foil. The X-rays are scattered by the foil and detected via a CCD (charge coupled device) behind the foil which forms a digital image of the resulting pattern.



a If the wavelength of the X-rays is 260 pm (260×10^{-12} m), what is the energy of the X-rays?

b The CCD displays an image on a computer screen of the diffraction pattern formed once the X-rays have passed through the foil. A beam of accelerated electrons is then substituted for the X-rays. A very similar diffraction pattern is observed. Why do the electrons produce a diffraction pattern with a similar spacing to that of the original X-rays?





c Based on the assumption that the two diffraction patterns have the same radius, what is the momentum of the electrons?

10 Davisson and Germer conducted an experiment where electrons were scattered after being fired at a target.

a What was observed by a detector moving through the scattering angles?

b What was the implication of this?

11 Which of the following would have the longest wavelength?

A electron, $m = 9.1 \times 10^{-31}$ kg, $v = 7.5 \times 10^{6}$ m s^{-1}

B blue light, $\lambda = 470$ nm

C X-ray, $f = 5 \times 10^{17}$ Hz

D proton, momentum = 1.7×10^{-21} kg m s^{-1}

12 What is the de Broglie wavelength of a 40 g bullet travelling at 1.0×10^{3} m s^{-1}?

13 Would wave behaviours such as diffraction be noticeable for the bullet described in Question 12?

14 A particular atom has four energy levels. In this context, what does it mean to say that the levels are *quantised*?

15 Calculate the frequency of the photon produced when an electron in a hydrogen atom drops from the $n = 3$ energy level to its ground state, $n = 1$. Refer to Figure 10.3.8.

16 Explain why the development of the Bohr model of the hydrogen atom was significant in the development of a comprehensive understanding of the nature of light.

17 Describe the relationship between the colours seen in the emission and absorption spectra of hydrogen.

18 By referring to the behaviour of electrons, explain how light is produced in an incandescent (filament) light globe.

19 According to Heisenberg's uncertainty principle, if the uncertainty in position is decreased what will happen to the uncertainty in momentum?

20 If a photon of a very short wavelength were to collide with an electron, what would be the effect on the position of the electron?

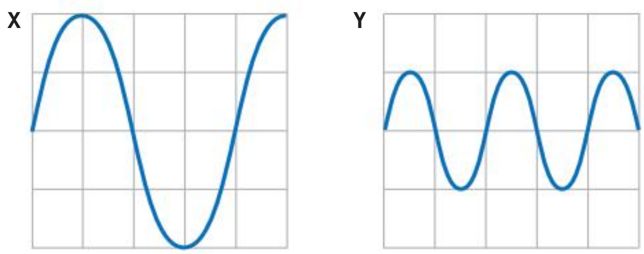
UNIT 4 • Area of Study 1

REVIEW QUESTIONS

How can waves explain the behaviour of light?

The following information relates to questions 1–6.

A microphone is placed in front of a sound source. The output is fed into a digitiser connected to a computer. X and Y are the traces from two different sounds shown on the screen of the computer. The sampling settings are identical for each trace.



1 Which of the following do the traces represent?

A Pressure variation versus time of a transverse wave.
B Displacement of molecules versus time of a transverse wave.
C Pressure variation versus time for a longitudinal wave.
D Movement of individual molecules in the air up and down directly representing the sound wave.

2 What is the ratio of the amplitudes $A_X : A_Y$?

3 What is the ratio of the frequencies $f_X : f_Y$?

4 How many rarefactions of wave X are represented in trace X?

5 How many compressions of wave Y are represented in trace Y?

6 Which one (or more) of the following statements about the air pressure in a sound wave as it relates to compressions and rarefactions is true?

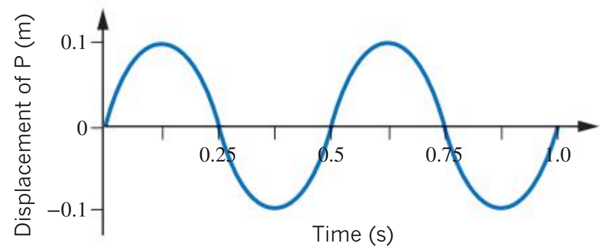
A Pressure is at a maximum halfway between a compression and a rarefaction.
B Pressure is at a minimum halfway between a compression and a rarefaction.
C Pressure is at a minimum at a compression and it is at a maximum at a rarefaction.
D Pressure is at a minimum at a rarefaction and it is at a maximum at a compression.

7 A student moves a slinky spring down and then up before returning it to the original position. A pulse moves to the right, away from the student's hand. Draw a diagram that represents the pulse formed in the spring by this action.

8 Define what is meant by a mechanical wave.

9 Explain the difference between a transverse wave and a longitudinal wave.

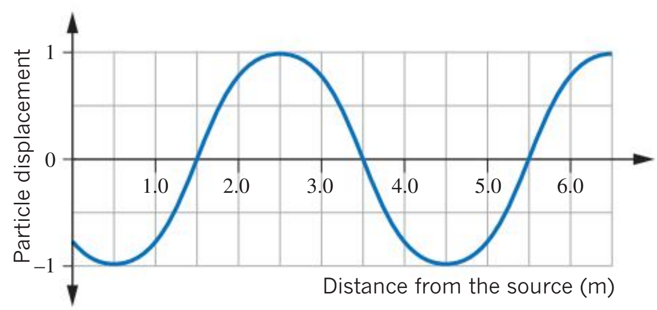
10 This is the displacement–time graph for a particle P.



Draw a graph that shows the displacement–time graph for a particle, Q, positioned $\frac{3}{4}$ of a wavelength ahead of particle P.

The following information relates to questions 11 and 12.

The diagram shows the displacement of the air molecules in a sound wave from their mean positions as a function of distance from the source, at a particular time. The wave is travelling to the right at 340 m s^{-1}.



11 What is the wavelength of the sound wave?

12 Which arrow below describes the direction of transfer of acoustic energy by this wave?

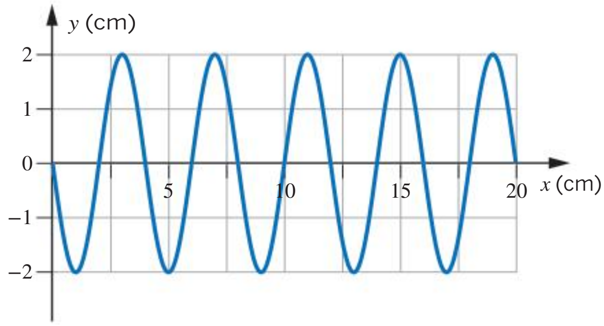
A →
B ←
C ↑
D ↓
E no energy is transferred

13 Which of the following properties of sound is independent of the source producing the sound?

A frequency
B amplitude
C speed

14 If a sound wave has a period of 3.9×10^{-3} s and the speed of sound is 340 m s^{-1}, calculate the wavelength of this sound wave.

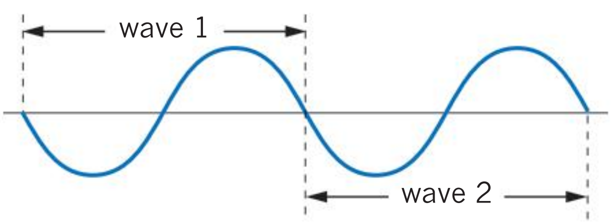
15 Determine the wavelength and amplitude of the wave depicted in the following graph.



16 A wave travels along a rope until it reaches the fixed end of the rope. Describe fully what would occur after this point in time.

17 Describe what happens when two or more waves meet while travelling through the same medium.

18 Two identical wave pulses moving in opposite directions towards each other, meet as shown in the diagram below. Once each wave has moved along by another half a wavelength, describe the amplitude and wavelength of the wave that results due to superposition.



19 Two wave pulses along a string interact. As the two pulses pass through each other, is there any permanent alteration to either pulse? Explain your answer.

The following information relates to questions 20–22.

A string is attached to a ring around a pole so that it is free to move. The other end of the string is fixed. The tension in the string is constant. The effective length of the string is 75 cm and the speed of a wave created in the string is 330 m s^{-1}.

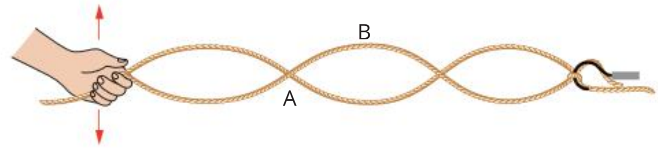
20 Calculate the fundamental frequency (also known as the first harmonic).

21 Calculate the frequency of the third harmonic.

22 Calculate the frequency of the next harmonic after the third that the string can produce.

The following information relates to questions 23–25.

The following diagram shows the standing wave pattern made by a rope when it is oscillated back and forth.

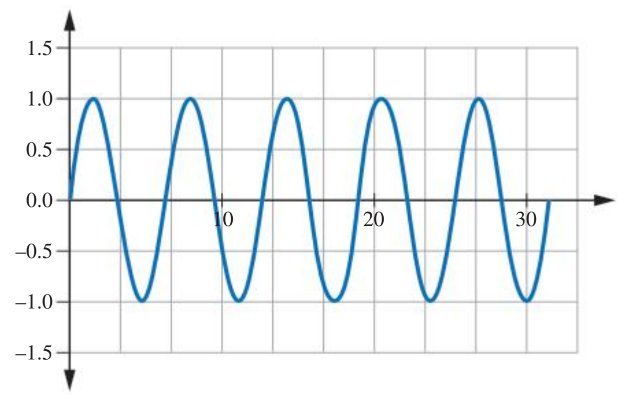


23 What are the names given to points A and B?

24 Why does the image show two ropes?

25 Which harmonic is shown in the image?

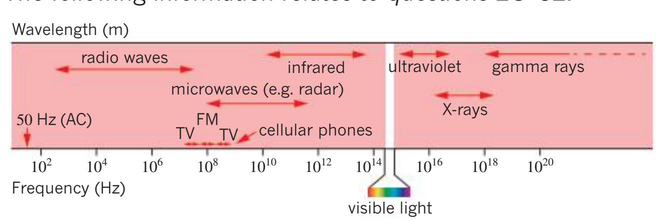
26 The diagram below shows the wave form from a microwave source, when the source is stationary with respect to the detector.



What will happen to the wave form when the microwave source is moving towards the stationary detector?

27 How does resonance apply in the case of pushing someone on a swing?

The following information relates to questions 28–32.



28 At what speed do the waves shown in the diagram travel in air?

29 What is the wavelength of waves with a frequency of 10^{16} Hz?

30 What specific type of waves are those described in Question 29?

31 Explain one helpful use of the waves described in Question 29.

32 A 60 W lamp emits radiation of wavelength 3.0×10^{-5} m. What type of light is the lamp emitting?

33 Sound waves of frequency *f* are diffracted as they pass through a narrow slit of width *w*. The amount of diffraction can be increased by (choose one or more answers):

A increasing *f*
B increasing *w*
C decreasing *f*
D decreasing *w*

The following information relates to questions 34 and 35.
Light travelling in air enters a second medium of refractive index 2.42 at an angle of incidence of 30.0°.

34 What is the angle of refraction?

35 Calculate the speed of light in medium 2.

36 How do the different colours of visible light differ from each other?

The following information relates to questions 37 and 38.
Two students were given the task by their teacher to demonstrate the dispersion of white light in the school science laboratory.

37 What equipment would they need and how would they demonstrate dispersion?

38 What result will the students achieve if they perform the experiment correctly?

39 What does it mean if light is 'polarised'?

40 Explain why it is impossible to polarise sound waves.

The following information relates to questions 41 and 42.
Two polarising sheets are placed together, with their polarising axes parallel, and are held up to a light source. One of the polarising sheets is rotated through 360° while the other is held still. The transmitted light alternates between maximum and zero intensities for different alignments of the polarising sheets.

41 What alignments of the two polarising sheets produce the maximum light intensities?

42 Why are there two maximums and two zeros of intensity during the rotations?

The following information refers to questions 43–46.
A Young's double-slit experiment is performed using red light and the resulting interference pattern is observed.

43 What is the effect on the observed interference pattern of halving the separation of the slits?

44 What is the effect on the observed interference pattern of doubling the distance between the slits and the screen?

45 What is the effect on the observed interference pattern of halving the frequency of the light used?

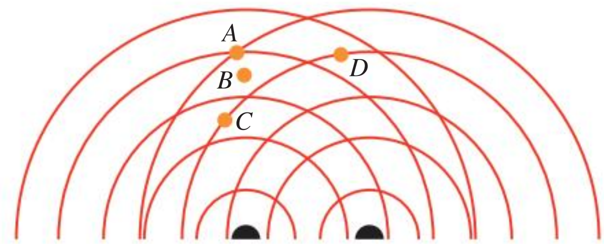
46 What is the effect on the observed interference pattern of covering one of the slits?

The following information refers to questions 47 and 48.
At the time when Thomas Young carried out his famous double-slit experiment, there were two competing models claiming to explain the nature of light.

47 What were the names of the two competing models?

48 Explain how Young's experiment supported one of these models and not the other.

49 The image below shows the wave fronts from two sources of waves. The red lines represent wave crests.



Identify the positions where maximum constructive interference occurs. Explain your answer.

50 Physicists replicating Young's famous double-slit experiment determine that particular adjacent dark bands on the interference pattern (e.g. the third dark fringe and the fourth dark fringe) are different in their distance from one of the slits by only 500 nm. Determine the wavelength of the monochromatic light being used.

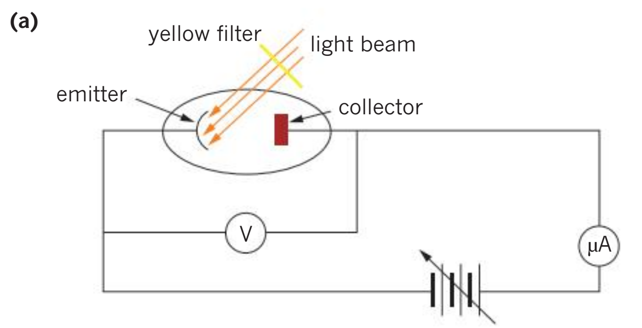
UNIT 4 • Area of Study 2

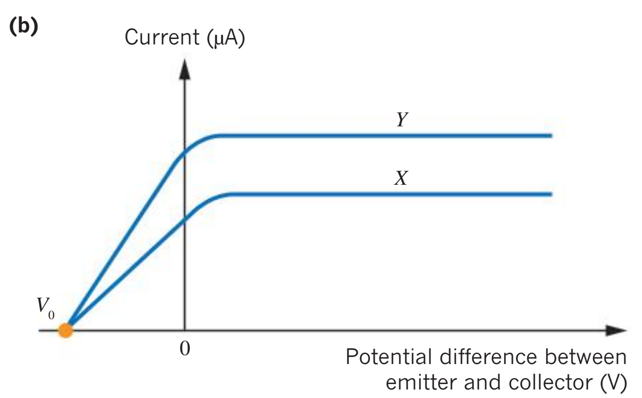
REVIEW QUESTIONS

How are light and matter similar?

The following information relates to questions 1–4.

Light passing through a yellow filter is incident on the cathode in a photoelectric effect experiment as shown in diagram (a). The reverse current in the circuit can be altered using a variable voltage. At the stopping voltage, V_0, the photocurrent is zero. The current in the circuit is plotted as a function of the applied voltage in diagram (b).





1 Which of the following changes would result in an increase in the size of V_0?

A replacing the yellow filter with a red filter

B replacing the yellow filter with a blue filter

C increasing the intensity of the yellow light

2 Which one of the following options best describes why there is zero current in the circuit when the applied voltage equals the stopping voltage?

A The threshold frequency of the emitter increases to a value higher than the frequency of yellow light.

B The work function of the emitter is increased to a value higher than the energy of a photon of yellow light.

C The emitted photoelectrons do not have enough kinetic energy to reach the collector.

3 Which of the following descriptions of the graphs X and Y in diagram (b) are correct?

A Both graphs are produced by yellow light of different intensities.

B Graph X is produced by yellow light while graph Y is produced by blue light.

C Each graph is produced by light of a different colour and different intensity.

4 The emitter of the photocell is coated with nickel. The filter is removed and a 200 nm light is directed onto the cathode. The minimum value of V_0 that will result in zero current in the circuit is 1.21 V. What is the work function of nickel?

5 Describe three experimental results associated with the photoelectric effect that cannot be explained by the wave model of light.

The following information relates to questions 6–9.

In a double-slit interference experiment, an electron beam travels through two narrow slits, 20 mm apart, in a piece of copper foil. The resulting pattern is detected photographically at a distance of 2.0 m. The speed of the electrons is 0.1% of the speed of light.

6 Calculate the de Broglie wavelength of the electrons used in the experiment.

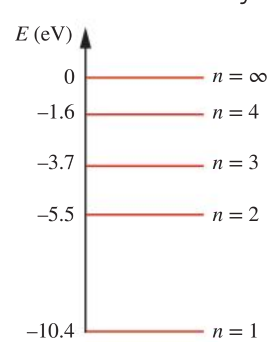
7 What do you expect to see on the photographic plate?

8 Given that electrons are particles, how do you interpret the behaviour of the electrons in this experiment?

9 If the experiment were to be repeated using neutrons, at what speed would a neutron need to travel to have the same de Broglie wavelength as the electrons in Question 8?

The following information relates to questions 10–12.

The energy levels for atomic mercury are as follows.



Determine the frequency and wavelength of the light emitted when the atom makes the following transitions:

10 $n = 4$ to $n = 1$

11 $n = 2$ to $n = 1$

12 $n = 4$ to $n = 3$

The following information relates to questions 13–15.
An electron is accelerated across a potential difference of 65 V.

13 What kinetic energy will the electron gain?

14 What speed will the electron reach?

15 What is the de Broglie wavelength of the electron?

16 How did Niels Bohr explain the observation that for the hydrogen atom, when the frequency of incident light was below a certain value, the light would simply pass straight through a sample of hydrogen gas without any absorption occurring?

The following information relates to questions 17–19.
Physicists can investigate the spacing of atoms in a powdered crystal sample using electron diffraction. This involves accelerating electrons to known speeds using an accelerating voltage. In a particular experiment, electrons of mass 9.11×10^{-31} kg are accelerated to a speed of 1.75×10^{7} m s^{-1}. The electrons pass through a powdered crystal sample, and the diffraction pattern is observed on a fluorescent screen.

17 Calculate the De Broglie wavelength (in nm) of the accelerated electrons.

18 Describe the main features of the expected diffraction pattern.

19 If the accelerating voltage is increased, what difference would you expect to see in the diffraction pattern produced? Explain your answer.

20 How would de Broglie explain the light and dark rings produced when a beam of electrons is fired through a sodium chloride crystal?

21 Describe how the wave–particle duality of electrons can be used to explain the quantised energy levels of the atoms.

22 Which one or more of the following phenomena can be modelled by a pure wave model of light?
- **A** the photoelectric effect
- **B** refraction
- **C** the double-slit interference of light
- **D** reflection
- **E** diffraction

23 Define the electron-volt.

24 Why are all of the frequencies of light above the ionisation energy value for hydrogen continuously absorbed?

25 How do our wave and particle models of light parallel the ideas related to electrons and matter waves?

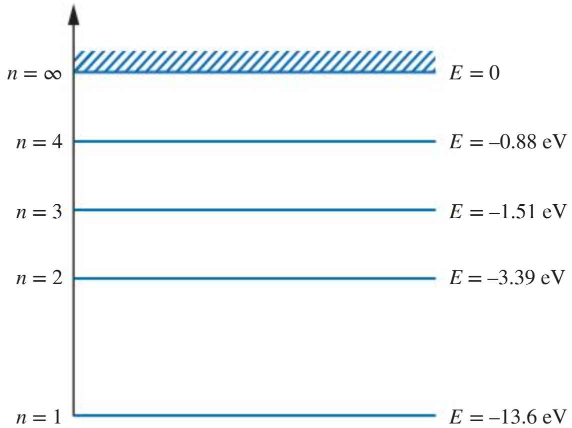
26 For an electron and a proton to have the same wavelength:
- **A** the electron must have the same energy as the proton.
- **B** the electron must have the same speed as the proton.
- **C** the electron must have the same momentum as the proton.
- **D** It is impossible for an electron and a proton to have the same wavelength.

The following information relates to questions 27 and 28.
When conducting a photoelectric effect experiment, a student correctly observes that the energy of emitted electrons depended only on the frequency of the incident light and was independent of the intensity.

27 Explain how the particle model accounts for this observation.

28 Explain why the wave model cannot account for this observation.

The following information relates to questions 29–33.
Consider the energy-level diagram for the hydrogen atom shown below. A photon of energy 14.0 eV collided with a hydrogen atom in the ground state.



29 Explain why this collision will eject an electron from the atom.

30 Calculate the energy of the ejected electron in electronvolts and in joules.

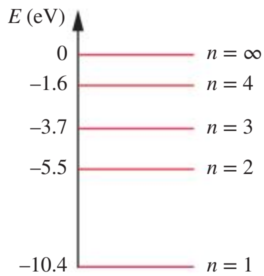
31 What is the momentum of the ejected electron?

32 Determine the wavelength of the ejected electron.

33 A hydrogen atom in the ground state collides with a 10.0 eV photon. Describe the result of such a collision.

The following information relates to questions 34–36.

An electron beam of energy 7.0 eV passes through some mercury vapour in the ground state and excites the electrons to the $n = 3$ energy level.



34 List all the possible photon energies that would be present in the emission spectrum when the electrons return from the $n = 3$ energy level to the ground state.

35 What is the shortest wavelength of light present in the emission spectrum?

36 A photon collides with a mercury atom in the ground state. As a result, a 30.4 eV electron is ejected from the atom. What was the wavelength of the incident photon?

The following information relates to questions 37–39.

When investigating the photoelectric effect, the relationship between the maximum kinetic energy of emitted photoelectrons and the frequency of the light incident on the metal plate is:

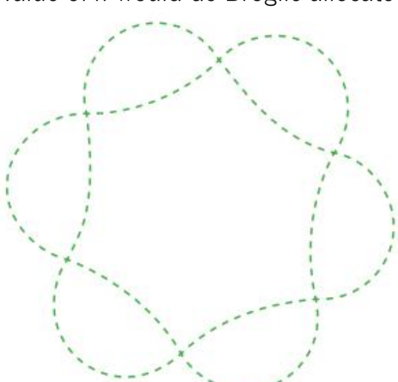
$E_{k\ max} = hf - \phi$

37 Explain the meaning of the terms $E_{k\ max}$, f and ϕ in this equation.

38 If the intensity of the light striking the metal is increased, but the frequency is unaltered, what effect does this have on the value of $E_{k\ max}$?

39 If the intensity of the light striking the metal is increased, but the frequency is unaltered, what effect does this have on the value of the current flowing in the apparatus?

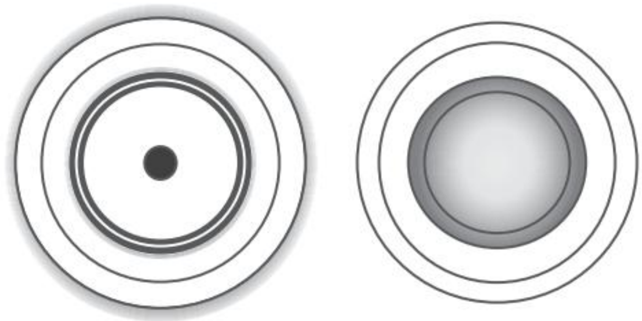
40 The diagram below represents the 'standing-wave state' of an electron in an atom of hydrogen. Which value of n would de Broglie allocate to this pattern?



41 What is the momentum of a gamma ray with a wavelength of 3.0 pm?

The following information relates to questions 42–46.

The image below shows diffraction images that have been obtained by scattering (a) X-rays and (b) electrons off the same sample, which is made up of many tiny crystals with random orientation. The X-rays have a frequency of 8.3×10^{18} Hz.



42 Provide an explanation for the fact that the electrons and the X-rays have produced the same diffraction pattern.

43 Determine the wavelength of the X-ray photons.

44 Determine the de Broglie wavelength of the electrons.

45 Calculate the momentum of the electrons.

46 Do the X-rays and the electrons have the same energy? Explain your answer.

47 Make three statements about how the particle (photon) model of light is supported by features of the photoelectric effect and discuss the implications for the wave model of light in each case.

48 The light of which of the options A–D is:

a generated by the electrons in individual atoms as they drop one or more energy levels?

b generated by random thermal motion of atoms in the material?

c generated as electrons fall from the conduction band to the valence band in a semiconductor?

d coherent

A incandescent light bulb

B sodium vapour lamp

C light-emitting diode

D laser

49 Explain how the single-slit diffraction experiment relates to the Heisenberg uncertainty principle.

50 The Heisenberg uncertainty principle states:

$\Delta p \Delta x \geq \frac{h}{4\pi}$

What does it mean if the value of Δp in the relation gets smaller? Explain your answer.

APPENDIX A SI units

THE STANDARD UNITS OF MEASUREMENT

The accurate and easy measurement of quantities is essential in both everyday life and for scientific investigation. Over the centuries, many different systems of measuring physical quantities have been developed. For example, length can be measured in chains, fathoms, furlongs, yards, feet, rods and microns. Some units were based on parts of the body. The cubit was defined as the distance from the elbow to the fingertip, and so the amount of cloth that you obtained from a tailor depended on the physical size of the person selling it to you.

The metric system was established by the French Academy of Science at the time of the French Revolution (1789–1815) and is now used in most countries. This system includes units such as the metre, litre and kilogram. Countries of the British Empire adopted the British Imperial system of the mile, gallon and pound. These two systems developed independently and their dual existence created problems in areas such as trade and scientific research. In 1960, an international committee set standard units for fundamental physical quantities. This system was an adaptation of the metric system and is known as the *Système Internationale d'Unités* (International System of Units) or SI system of units.

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

TABLE A.1 The SI units identify the seven fundamental quantities whose basic value is defined to a high degree of accuracy.

Mass

The kilogram was originally defined as the mass of 1 L of water at 4°C. This is still approximately correct, but a far more precise definition is now used. Since 1897 the measurement standard for the kilogram has been a cylindrical block of platinum–iridium alloy kept at the International Bureau of Weights and Measures in France. Australia has a copy of this standard mass at the CSIRO Division of Applied Physics in Sydney. At times it is returned to France to ensure that the mass remains accurate.

Length

The metre was originally defined in 1792 as one ten-millionth of the distance from the equator to the North Pole (approximately 10 000 km). This definition has changed a number of times since. In 1983, to give a more accurate value, the metre was redefined as the distance that light in a vacuum travels in $\frac{1}{299\,792\,458}$ second. This standard can be reproduced all over the world, as light travels at a constant speed in a vacuum.

PHYSICSFILE

Metric system

The metric system was originally developed in France and is known as the *Système Internationale* (SI). It was adopted in France in 1840 as the official system of units, although it had been developing in that country since 1545. It has remained in use ever since and has gradually been adopted by most other countries. It has been modified a little over the years and now, in Australia, we use SI units that have been standardised by the International Standards Organisation (ISO) since the 1960s. Some countries such as France, Italy and Spain use an earlier form of the metric system that is slightly different. The USA still measures almost everything in the old imperial units such as pounds for mass and feet for distance but, even there, scientists use the SI system of units. There are two major advantages of using the metric system. It is easier to use than other systems in that derived units are straightforward and various sizes of units are created using multiples of ten. The other very big advantage is the international nature of the standards and units. All units are standardised, making comparisons straightforward.

Time

Up to 1967, time had always been based on the apparent motion of the heavens. The second was once defined in terms of the motion of the Sun. Until 1960, one second was defined as $\frac{1}{60}$ of $\frac{1}{60}$ of $\frac{1}{24}$ of an average day in 1900. This reflected the rate of the Earth's rotation on its axis; however, its rotation is not quite uniform. In 1967, a more accurate definition was adopted—one not based on the motion of the Earth. One second is now defined as the time required for a caesium-133 atom to undergo 9 162 631 770 vibrations. These vibrations are stimulated by an electric current and are extremely stable, allowing this standard to be reproduced all over the world.

DERIVED UNITS

As well as the seven fundamental quantities, a wide variety of other physical quantities can be measured. You may have encountered some of these, such as frequency, velocity, energy and density, already. A derived quantity is defined in terms of the fundamental quantities. For example, the SI unit for area is square metres (m^2).

Quantity	SI unit	SI unit symbol	Equivalent unit
velocity	metres per second	$m\ s^{-1}$	—
acceleration	metres per second per second	$m\ s^{-2}$	—
frequency	hertz	Hz	s^{-1}
force	newton	N	$kg\ m\ s^{-2}$
energy/work	joule	J	$kg\ m^2\ s^{-2}$

TABLE A.2 Some derived SI quantities and their units.

MEASUREMENT AND UNITS

In every area of physics we have attempted to quantify the phenomena we study. In practical demonstrations and investigations we generally make measurements and process those measurements in order to come to some conclusions. Scientists have a number of conventional ways of interpreting and analysing data from their investigations. There are also conventional ways of writing numerical measurements and their units.

Correct use of unit symbols

The correct use of unit symbols removes ambiguity, as symbols are recognised internationally. The symbols for units are not abbreviations and should not be followed by a full stop unless they are at the end of a sentence.

Upper-case letters are not used for the names of any physical quantities of units. For example, we write newton for the unit of force, while we write Newton if referring to someone with that name. Upper-case letters are only used for the *symbols* of the units that are named after people. For example, the unit of energy is joule and the symbol is J. The joule was named after James Joule who was famous for studies into energy conversions. The exception to this rule is 'L' for litre. We do this because a lower-case 'l' looks like the numeral '1'. The unit of distance is metre and the symbol is m. The metre is not named after a person.

The product of a number of units is shown by separating the symbol for each unit with a dot or a space. Most teachers prefer a space but a dot is perfectly correct. The division or ratio of two or more units can be shown in fraction form, using a slash, or using negative indices. Most teachers prefer negative indices. Prefixes should not be separated by a space.

Preferred	Correct also		Wrong	
$m\ s^{-2}$	$m.s^{-2}$	m/s^2	ms^{-2}	
kW h	kW.h		kWh	k Wh
$kg\ m^{-3}$	$kg.m^{-3}$	kg/m^3	kgm^{-3}	
μm			μ m	
N m	N.m		Nm	

TABLE B.1 Some examples of the use of symbols for derived units.

Units named after people can take the plural form by adding an 's' when used with numbers greater than one. Never do this with the unit symbols. It is acceptable to say 'two newtons' but wrong to write 2 Ns. It is also acceptable to say 'two newton'.

Numbers and symbols should not be mixed with words for units and numbers. For example, twenty metres and 20 m are correct while 20 metres and twenty m are incorrect.

Scientific notation

To overcome confusion or ambiguity, measurements are often written in scientific notation. Quantities are written as a number between one and ten and then multiplied by an appropriate power of ten. Note that 'scientific notation', 'standard notation' and 'standard form' all have the same meaning.

Examples of some measurements written in scientific notation are:

$0.054\ m = 5.4 \times 10^{-2}\ m$

$245.7\ J = 2.457 \times 10^{2}\ J$

$2080\ N = 2.080 \times 10^{3}\ N$ or $2.08 \times 10^{3}\ N$

You should be routinely using scientific notation to express numbers. This also involves learning to use your calculator intelligently. Scientific and graphics calculators can be put into a mode whereby all numbers are displayed in scientific notation. It is useful when doing calculations to use this mode rather than frequently attempting to convert to scientific notation by counting digits on the calculator display. It is quite acceptable to write all numbers in scientific notation, although most people prefer not to use scientific notation when writing numbers between 0.1 and 1000.

FIGURE B.1 A scientific calculator.

An important reason for using scientific notation is that it removes ambiguity about the precision of some measurements. For example, a measurement recorded as 240 m could be a measurement to the nearest metre; that is, somewhere between 239.5 m and 240.5 m. It could also be a measurement to the nearest ten metres, that is, somewhere between 235 m and 245 m. Writing the measurement as 240 m does not indicate either case. If the measurement was taken to the nearest metre, it would be written in scientific notation as 2.40×10^2 m. If it was taken to the nearest ten metres only, it would be written as 2.4×10^2 m.

PREFIXES AND CONVERSION FACTORS

Conversion factors should be used carefully. You should be familiar with the prefixes and conversion factors in Table B.2. The most common mistake made with conversion factors is multiplying rather than dividing. Some simple strategies can save you this problem. Note that the table gives all conversions as a multiplying factor.

Multiplying factor		Prefix	Symbol
1 000 000 000 000	10^{12}	tera	T
1 000 000 000	10^{9}	giga	G
1 000 000	10^{6}	mega	M
1 000	10^{3}	kilo	k
0.01	10^{-2}	centi	c
0.001	10^{-3}	milli	m
0.000 001	10^{-6}	micro	µ
0.000 000 001	10^{-9}	nano	n
0.000 000 000 001	10^{-12}	pico	p

Do not put spaces between prefixes and unit symbols. It is important to give the symbol the correct case (upper or lower case). There is a big difference between 1 mm and 1 Mm.

TABLE B.2 Prefixes and conversion factors.

There is no space between prefixes and unit symbols. For example, one-thousandth of an ampere is given the symbol mA. Writing it as m A is incorrect. The space would mean that the symbol is for a derived unit—a metre ampere.

Worked example B1

The diameter of a cylindrical piece of copper rod was measured at 24.8 mm with a vernier caliper. Its length was measured at 35 cm with a tape measure.

a Find the area of cross-section in m^2.

b Find the volume of the copper rod in m^3.

Solution

a The area of cross-section is πr^2. The radius is calculated by dividing the diameter by two. Hence the radius is 12.4 mm. To calculate the area in m^2, first halve the diameter and convert it to metres. The radius is $\frac{24.8}{2}$ = 12.4 mm = 12.4×10^{-3} m. The radius is not written in scientific notation. This is not necessary. All you need to do is multiply by the appropriate factor. The conversion factor for mm to m is 10^{-3}. Just multiply by the conversion factor and don't bother to rewrite the result in scientific notation. This is because it is only going to be used in a calculation and is not a final result.

The area of cross-section is $\pi r^2 = \pi(12.4 \times 10^{-3})^2 = 4.8 \times 10^{-4}$ m^2.

b The volume is $\pi r^2 h$, where h is the length of the cylinder.

The length is 35 cm = 35×10^{-2} m.

Hence the volume is $\pi(12.4 \times 10^{-3})^2(35 \times 10^{-2}) = 1.7 \times 10^{-4}$ m^2.

Worked example B2

a A car is traveling at 110 km h^{-1}. How fast is this in m s^{-1}?

b Convert 35 miles per hour to metres per second. A mile is approximately 1600 m.

Solution

a 110 km h^{-1} is 110 × 10^3 metres per 3600 s.

$$\frac{110 \times 10^3}{3600} = 30.6$$

Hence 110 km h^{-1} = 30.6 m s^{-1}.

b 35 miles per hour is 35 × 1600 metres per 3600 s.

$$\frac{35 \times 1600}{3600} = 15.6$$

Hence 35 mph = 16 m s^{-1}.

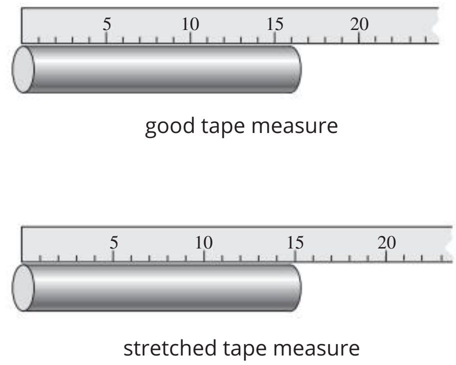


FIGURE B.2 The diagram shows that a correctly manufactured tape measure correctly measures the cylinder to be 16 cm long while the stretched tape measure gives a wrong measurement of 15 cm. The stretched tape measure is inaccurate.

DATA

Physicists and physics students collect, analyse and interpret experimental data. In fact, you will do this when you conduct your Practical Investigation in Unit 2 (Area of Study 3). Working with data requires a good understanding of the meaning and limitations of measurement.

Accuracy and precision

Two very important aspects of any measurement are accuracy and precision. Accuracy and precision are not the same thing. The distinction between the two ideas is only hard to grasp because the two words are defined in a similar way in the dictionary. We often hear the words used together and in general conversation they tend to be used interchangeably.

Instruments are said to be *accurate* if they truly reflect the quantity being measured. For example, if a tape measure is correctly manufactured it can be used to measure lengths accurately to the nearest centimetre.

Imagine that the tape measure is accidentally stretched during the manufacturing process, as shown in Figure B.2. It would still be used to measure length to the nearest centimetre but all measurements would be wrong. It would be inaccurate.

Suppose an accurate ruler had 3 cm snapped off the end, as shown in Figure B.3. It would now give readings all too large by 3 cm if no allowance were made for the missing piece. This ruler measure would be inaccurate.

In these two examples, the tape measure or ruler is used to measure to the nearest centimetre but is inaccurate. Inaccurate means just plain wrong. Instruments are said to be *precise* if they can differentiate between slightly different quantities. Precision refers to the fineness of the scale being used.

Consider the metre rule, the tape measure and the measuring wheel used to mark out sports fields. All three measure distance. All three can be accurate. The metre rule is more *precise* because it measures to the nearest millimetre, the tape measure has less precision due to measuring only to the nearest centimetre, while the wheel measures only to the nearest metre (Figure B.4).

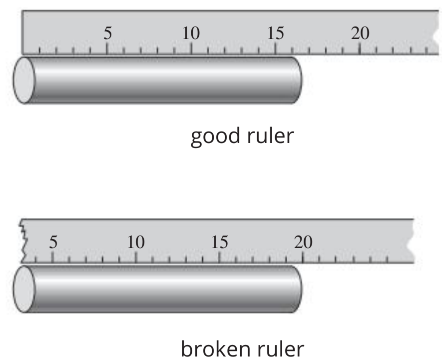


FIGURE B.3 The diagram shows that an undamaged ruler correctly measures the cylinder to be 16 cm long while the broken ruler gives a wrong measurement of 19 cm. The broken ruler is inaccurate but equally as precise as the unbroken ruler.

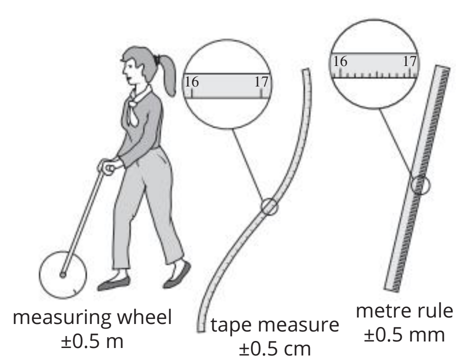


FIGURE B.4 The measuring wheel has low precision and only measures to the nearest metre. It has an uncertainty of 0.5 m. The tape measure has more precision and has an uncertainty of 0.5 cm or 0.005 m. The metre rule has an uncertainty of 0.5 mm or 0.0005 m.

The tape measure is a more precise instrument than the measuring wheel. Suppose two distances of 2673 mm and 2691 mm are being measured with these two instruments. Each distance would be measured as 3 m, to the nearest metre, by the wheel. They would be measured differently as 2.67 m and 2.69 m, to the nearest centimetre, by the tape measure. The tape measure is more precise because it has a finer scale. You might also say that it has greater resolution. The measuring wheel has such low precision that it can't be used to measure which of the two distances is greater or smaller. Measuring instruments with less precision give measurements that are less certain. The uncertainty in the measurement is due to a coarser scale. The measuring wheel gives less certain measurements than the tape measure even though both instruments may be equally accurate.

All measurements have some amount of *uncertainty*, due to the precision of the instrument which does the measuring. The uncertainty is generally one half of the finest scale division on the measuring instrument. The measuring wheel has an uncertainty of 0.5 m. The metre rule has an uncertainty of 0.5 mm. The tape measure has an uncertainty of 0.5 cm. An electronic balance set to measure grams to two decimal places has an uncertainty of 0.005 g.

Sometimes this uncertainty is referred to as error. It is not error, in that it is not a mistake or something wrong. All measuring instruments have limited precision and, in general, the uncertainty is half of the smallest scale division on the instrument.

The uncertainty is, indeed, the measure of the precision of an instrument. It is not related to accuracy. A micrometer screw gauge, which measures length to the nearest one-hundredth of a millimetre and hence is very precise, may not be accurate. Usually they are, but if one has been badly manufactured or bent by being over-tightened repeatedly it most likely will be inaccurate. But its precision will still be ±0.000 005 m, or half of one-hundredth of a millimetre.

The uncertainty gives the range in which a measurement falls. If you measured the length of a stick with a metre rule then you would get a measurement 'plus or minus' half a millimetre.

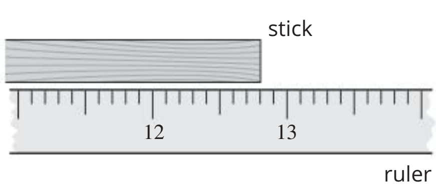


FIGURE B.5 A stick anywhere between 127.5 mm and 128.5 mm would be recorded as having a length of 128 mm if measured by a metre rule with a scale division of 1 mm. Conversely, a measurement recorded as 128 mm could be of an object of length anywhere between 127.5 mm and 128.5 mm.

Any stick between 127.5 mm and 128.5 mm long would be measured as 128 mm to the nearest millimetre (refer to Figure B.5). We would record this as 128 ± 0.5 mm.

When using an analogue scale, you might think that you can 'judge by eye' fractions of a scale division and hence get greater precision than half a scale division. You should be able to judge to the nearest half a scale division. You might think you can judge to the nearest tenth of a division. You can't. Research shows that despite the fact that people try to judge the spaces between scale divisions to better than half a division, as soon as this is done, inconsistent measurements are obtained. That is, different people get different measurements of the same thing.

The best judgement you can definitely claim is one half of a scale division. The uncertainty we will still assume, however, is a full half-scale division. Hence, you might measure another stick, one that has a length somewhere between 154 mm and 155 mm, as 154.5 ± 0.5 mm.

Of course, you don't have the option of adding an extra decimal place containing a 0 or a 5 if you are using a digital instrument.

The uncertainty can be recorded as the *absolute* uncertainty as we have done above. The absolute uncertainty is the actual uncertainty in the measurement. In this case it is 0.5 mm. Alternatively, it is often useful to write the uncertainty as a *percentage*: 0.5 mm is 0.32% of 154.5. Hence, the above length would be recorded as 154.5 mm ± 0.32%.

Percentage uncertainty is also called *relative* uncertainty. It is the size of the uncertainty relative to the size of the measured quantity.

PHYSICSFILE

Many people use the term 'error' to refer to uncertainty and many other things. The problem with referring to uncertainty as error is that it is not actually error. Things that are a normal consequence of the limitations of measuring instruments must happen, and are not mistakes. If they are not mistakes or 'something gone wrong' then it makes no sense to call them errors.

Errors are the factors that limit the accuracy of your results. For example, if you perform a calorimetry experiment and do not use a good enough insulator, you will get inaccurate results due to heat losses to the environment. This will contribute to the error in your measurement. Suppose you measured the refraction of light in glass but did not place the protractor in the correct place when measuring angles. This would also cause error.

Many different things can contribute to experimental error. Some are unavoidable. Some are factors in the design of experiments. Good experimental design seeks to eliminate or at least minimise potential sources of error.

Never quote 'human error' as a source of error. Your data should be examined carefully and mistakes eliminated or at least ignored. So-called human errors, or lack of care, have no place in your experimental work. If you make mistakes then you should repeat the measurements.

Estimating the uncertainty in a result

An experiment or a measurement exercise is not complete until the uncertainties have been analysed. It is also important to explain how uncertainty due to the precision of instruments affects results.

The following three processes are used for estimating uncertainty in calculations due to the precision of instruments. They are demonstrated in Worked example B3.

- When adding or subtracting data, add the absolute uncertainties.
- When multiplying or dividing data, add the percentage uncertainties.
- When raising data to power *n*, multiply the percentage uncertainty by *n*.

In Worked example B3, the analysis of uncertainty reveals the *precision* of an experimental result.

Worked example B3

In Area of Study 1 in Unit 1, you might have measured the specific heat capacity of a metal. You could have calculated your result using:

$$c_{metal} = \frac{c_{water} m_{water} \Delta T_{water}}{m_{metal} \Delta T_{metal}}$$

Suppose you had the following data included in your table.

Quantity		Absolute uncertainty	% uncertainty
c_{water}	4180 J kg^{-1} K^{-1}	5 J kg^{-1} K^{-1}	0.120
m_{water}	72.5×10^{-3} kg	0.05×10^{-3} kg	0.069
ΔT_{water}	5°C	1°C*	20
m_{metal}	87.3×10^{-3} kg	0.05×10^{-3} kg	0.057
ΔT_{metal}	72°C	1°C*	1.389

*Note that the ΔT values have an absolute uncertainty of 1°C because they are calculated by subtracting one temperature measurement from another.

PHYSICSFILE

In some classes, students are instructed to quote all results to two decimal places or to three significant figures. You should be able to see from Worked example B3 that these rules are not absolutely correct when applied to real data. For ordinary calculations in assignments, tests and examinations, you might just give your answers to three figures.

If a calculation is done in several stages then you should not round off any intermediate results. This will add rounding error to your calculations. Use the memory on your calculator so that there is no rounding until the end of your calculation.

You would calculate as follows:

$$c_{metal} = 241 \text{ J kg}^{-1} \text{ K}^{-1}$$

$$\text{Uncertainty (\%)} = 0.120 + 0.069 + 20 + 0.057 + 1.389 = 21.6\%$$

Hence, you would obtain the following result:

$$c_{metal} = 241 \text{ J kg}^{-1} \text{ K}^{-1} \pm 21.6\%$$

$$c_{metal} = 241 \pm 52 \text{ J kg}^{-1} \text{ K}^{-1}$$

Once you have done all of this you can consider the relative success of your measurement exercise.

Your result is:

$$189 \text{ J kg}^{-1} \text{ K}^{-1} \le c_{metal} \le 293 \text{ J kg}^{-1} \text{ K}^{-1}$$

If measurements by other people, such as the constants published in data books, fall within this range then you can conclude that your experiment is consistent with established values. That is, within the precision of your technique, there are probably no significant errors although the final measurement is rather imprecise in this case. We might say that it is accurate within the limitations of the equipment.

You are also now in a position to refine the experiment by reducing the larger uncertainties. In this case, the largest uncertainty was in the temperature change for the water. Hence, it would not be very helpful to measure the masses to greater precision because the limit to precision in this activity would be the temperature differences. Getting greater precision in the temperature changes would be a useful refinement.

You could consider ways of getting larger temperature changes in the water and hence obtain a smaller percentage uncertainty in the temperature change. Alternatively, you might consider ways of measuring the temperatures to greater precision.

If your measurement range does not include the result you expect, you should think about the origin of the errors. In other words, if you are sure that c_{metal} is less than 189 J kg^{-1} K^{-1} or more than 293 J kg^{-1} K^{-1} then there must be some error in your experimental technique or more uncertainty than you realised.

When reviewing an experiment or a measurement exercise, it is a good idea to consider both errors *and* uncertainties.

Significant figures

The number of significant figures in a measurement is simply the number of digits used when the number is written in scientific notation. Once you have done a calculation, your calculator usually has eight or ten digits in the display but most of them are meaningless. You must round off your answer appropriately.

Consider the result of the experiment described in Worked example B3. It would make no sense to quote the result to two decimal places (or five significant figures) when clearly the precision of the experiment gives less than three significant figures.

Calculated results never have more significant figures than the original data and might have fewer than the original data. If you are not doing a full analysis of the uncertainties, it is customary to give your answers to the same number of significant figures as the least precise piece of data. For example, in Worked example B3, the least precise data is the change in temperature of the water with only a single digit. The value for the specific heat might then be quoted simply as 2×10^2 J kg^{-1} K^{-1}, but doing the full calculation of the uncertainty in the result is much more informative.

GRAPHICAL ANALYSIS OF DATA

A major problem with doing a calculation from just one set of measurements is that a single incorrect measurement can significantly affect the result. Scientists like to take a large amount of data and observe the trends in that data. This gives more precise measurements and allows scientists to recognise and eliminate problematic data.

Physicists commonly use graphical techniques to analyse a set of data. In this section, the basic techniques that they use will be outlined and a general method for using a set of data that fits a known mathematical relationship will be developed.

Linear relationships

Some relationships studied in physics are linear, that is a straight line, while others are not. It is possible to manipulate non-linear data so that a linear graph reveals a measurement. Linear relationships and their graphs are fully specified with just two numbers: gradient, m, and vertical axis intercept, c. In general, linear relationships are written:

$$y = mx + c$$

The gradient, m, can be calculated from the coordinates of two points on the line:

$$m = \frac{\text{rise}}{\text{run}}$$

$$= \frac{y_2 - y_1}{x_2 - x_1}$$

where (x_1, y_1) and (x_2, y_2) are any two points on the line. Don't forget that m and c have units. Omitting these is a common error.

PHYSICSFILE

Graphs

When analysing data from a linear relationship, it is first necessary to obtain a graph of the data and an equation for the line that best fits the data. This line of best fit is often called the regression line. The entire process can be done on paper but most people will use a computer spreadsheet, the capabilities of a scientific or a graphics calculator, or some other computer-based process. In what follows, it is not assumed that you are using any particular technology.

If you are plotting your graph manually on paper then proceed as follows:

1. Plot each data point on clearly labelled, unbroken axes.
2. Identify and label but otherwise ignore any suspect data points.
3. Draw, by eye, the 'line of best fit' for the points. The points should be evenly scattered either side of the line.
4. Locate the vertical axis intercept and record its value as 'c'.
5. Choose two points on the line of best fit to calculate the gradient. Do not use two of the original data points as this will not give you the gradient of the line of best fit.
6. Write $y = mx + c$, replacing x and y with appropriate symbols, and use this equation for any further analysis.

If you are using a computer or a graphics calculator then proceed as follows:

1. Plot each data point on clearly labelled, unbroken axes.
2. Identify suspect data points and create another data table without the suspect data.
3. Plot a new graph without the suspect data. Keep both graphs as you don't actually discard the suspect data but do eliminate it from the analysis.
4. Plot the line of best fit—the regression line. The manner in which you do this depends on the model of calculator or the software being used.
5. Compute the equation of the line of best fit that will give you values for m and c.
6. Write $y = mx + c$, replacing x and y with appropriate symbols, and use this equation for any further analysis.

Worked example B4

Some students used a computer with an ultrasonic detector to obtain the speed–time data for a falling tennis ball. They wished to measure the acceleration of the ball as it fell. They assumed that the acceleration was nearly constant and that the relevant relationship was $v = u + at$, where v is the speed of the ball at any given time, u was the speed when the measurements began, a is the acceleration of the ball and t is the time since the measurement began.

Their computer returned the following data:

Time (s)	Speed (m s^{-1})
0.0	1.25
0.1	2.30
0.2	3.15
0.3	4.10
0.4	5.25
0.5	6.10
0.6	6.95

Find their experimental value for acceleration.

Solution

The data is assumed linear, with the relationship $v = u + at$, which can be thought of as being $v = at + u$, which makes it clear that putting v on the vertical axis and t on the horizontal axis gives a linear graph with gradient a and vertical intercept u. A graph of the data is shown in Figure B.6.

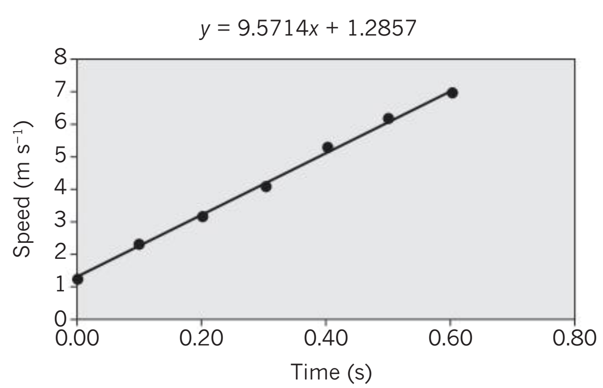


FIGURE B.6 Speed–time profile for a falling tennis ball.

This graph of the data was created on a computer spreadsheet. The line of best fit was created mathematically and plotted. The computer calculated the equation of the line. Graphics calculators can also do this.

A scientific calculator or graphics calculator or spreadsheet gives the regression line as $y = 9.5714x + 1.2857$. If this is rearranged and the constants are suitably rounded, the equation is $v = 1.3 + 9.6t$. This indicates that the ball was moving at 1.3 m s^{-1} at the commencement of data collection and the ball was accelerating at 9.6 m s^{-2}.

Manipulating non-linear data

Suppose you were examining the relationship between two quantities B and d and had good reason to believe that the relationship between them is

$$B = \frac{k}{d}$$

where k is some constant value. Clearly, this relationship is non-linear and a graph of B against d will not be a straight line. By thinking about the relationship it can be seen that in 'linear form':

$$\begin{array}{ccccc} B & = & k & \frac{1}{d} & \\ \uparrow & & \uparrow & \uparrow & \\ y & = & m & x & + c \end{array}$$

A graph of B (on the vertical axis) against $\frac{1}{d}$ (on the horizontal axis) will be linear. The gradient of the line will be k and the vertical intercept, c, will be zero. The line of best fit would be expected to go through the origin because, in this case, there is no constant added and so c is zero.

In the above example, a graph of the raw data would just show that B is larger as d is smaller. It would be impossible to determine the mathematical relationship just by looking at a graph of the raw data.

A graph of raw data will not give the mathematical relationship between the variables but can give some clues. The shape of the graph of raw data may suggest a possible relationship. Several relationships may be tried and then the best is chosen. Once this is done, it is not proof of the relationship but, possibly, strong evidence.

When an experiment involves a non-linear relationship, the following procedure is followed:

1. Plot a graph of the original raw data.
2. Choose a possible relationship based on the shape of the initial graph and your knowledge of various mathematical and graphical forms.
3. Work out how the data must be manipulated to give a linear graph.
4. Create a new data table.

Then follow the steps given in the Physics file on page 384. It may be necessary to try several mathematical forms to find one that seems to fit the data.

Worked example B5

Some students were investigating the relationship between current and resistance for a new solid-state electronic device. They obtained the data shown in the table.

According to the theory they had researched on relevant Internet sites, the students believed that the relationship between I and R is $R = dI^3 + g$, where d and g are constants.

By appropriate manipulation and graphical techniques, find their experimental values for d and g. The following steps should be used:

a Plot a graph of the raw data.

b Work out what you would have to graph to get a straight line.

c Make a new table of the manipulated data.

d Plot the graph of manipulated data.

e Find the equation relating I and R.

Current, I (A)	Resistance, R (Ω)
1.5	22
1.7	39
2.2	46
2.6	70
3.1	110
3.4	145
3.9	212
4.2	236

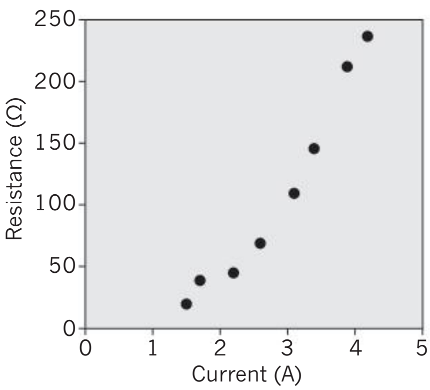


FIGURE B.7 Current–resistance graph of device.

Solution

a Figure B.7 shows the graph obtained using a spreadsheet.

It might be argued that the second piece of data is suspect. The rest of this solution supposes the students chose to ignore this piece of data.

b You can see what to graph if you think of the equation like this:

$$\begin{array}{ccccccc} R & = & d & I^3 & + & g \\ \uparrow & & \uparrow & \uparrow & & \uparrow \\ y & = & m & x & + & c \end{array}$$

A graph of R on the vertical axis and I^3 on the horizontal axis would have a gradient equal to d and a vertical axis intercept equal to g.

c The data is manipulated by finding the cube of each of the values for current.

Current cubed, I^3 (A^3)	Resistance, R (Ω)
3.38	22
10.65	46
17.58	70
29.79	110
39.30	145
59.32	212
74.09	236

d The graph in Figure B.8 was obtained from the spreadsheet.

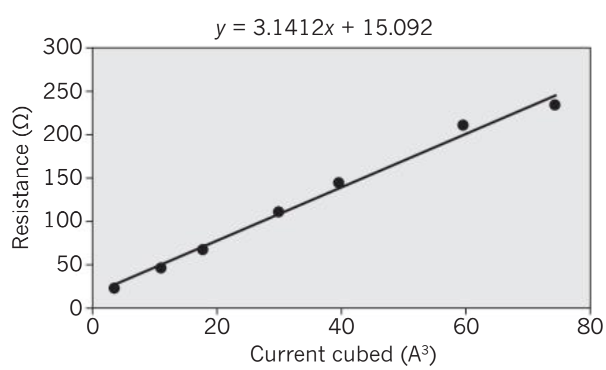


FIGURE B.8 Current–resistance characteristic (manipulated data).

e The regression line has the equation $y = 3.1x + 15.1$, so the equation relating I and R is $R = 3.1I^3 + 15.1$. Hence, the value of d is 3.1 Ω A^{-3} and the value of g is 15.1 Ω.

Answers

Chapter 1 Gravity

1.1 Newton's law of universal gravitation

WE 1.1.1 7.1×10^{-9} N **WE 1.1.2** 2.0×10^{20} N
WE 1.1.3 The acceleration of the Earth is 3.3×10^{5} times greater than the acceleration of the Sun.
WE 1.1.4 Both equations give the same result to two significant figures.
WE 1.1.5 Apparent weight = 810 N

1.1 review

1 The force of attraction between any two bodies in the universe is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.
2 the distance between the centres of the two objects
3 1.8×10^{21} N **4** 2.8×10^{-3} m s^{-2}
5 **a** 3.0×10^{16} N **b** 3.4×10^{22} N **c** 0.000088%
6 The Moon has a smaller mass than the Earth.
7 3.5 m s^{-2} **8** 240 N
9 Near the Earth's surface, weight is the gravitational force acting on an object whereas apparent weight is the contact force between the object and a surface. In an elevator accelerating upwards, a person's apparent weight would be greater than their weight.
10 **a** 550 N **b** 490 N

1.2 Gravitational fields

WE 1.2.1 **a** See Productlink **b** From strongest to weakest: B, C, A.
WE 1.2.2 9.6 N kg^{-1} **WE 1.2.3** 9.75 N kg^{-1} **WE 1.2.4** 3.7 N kg^{-1}

1.2 review

1 N kg^{-1} **2** 9.3 N kg^{-1} **3** $\frac{1}{9}$of the original
4 **a** 5.67 N kg^{-1} **b** 1.48 N kg^{-1}
c 0.56 N kg^{-1} **d** 0.22 N kg^{-1}
5 0.0008 N kg^{-1} or 8×10^{-4} N kg^{-1}
6 Mercury: 3.7 N kg^{-1}, Saturn: 10.4 N kg^{-1} Jupiter: 24.8 N kg^{-1}
7 2×10^{12} N kg^{-1}
8 The gravitational field strength at the poles is 1.4 times that at the equator.
9 3.4×10^{8} m **10** 10 Earth radii.

1.3 Work in a gravitational field

WE 1.3.1 3.1×10^{3} MJ **WE 1.3.2** 5.2 m s^{-1}
WE 1.3.3 5.2×10^{8} J **WE 1.3.4** 5.4×10^{9} J

1.3 review

1 C, hence its gravitational potential energy does not change. Its speed will also remain the same in a stable orbit.
2 It increases. **3** It will accelerate at an increasing rate.
4 A, B and C **5** 2.0×10^{12} J **6** 292 m s^{-1}
7 **a** between 9 N and 9.2 N **b** 2.6×10^{6} m or 2600 km
8 **a** 8×10^{6} J **b** 1.9×10^{7} J **c** 2.7×10^{7} J
d 7348 m s^{-1} or 7.3 km s^{-1}
9 1.7×10^{9} J **10** 2.6×10^{11} J

Chapter review

1 730 N **2** 3.78×10^{8} m **3** 2.1×10^{-7} m s^{-2}
4 **a** They are equal.
b The acceleration of Jupiter caused by the Sun is greater than the acceleration of the Sun caused by Jupiter.
5 3.7 m s^{-2} **6** **a** 460 N **b** 490 N
7 **a** 2.48×10^{4} N **b** 2.48×10^{4} N **c** 24.8 m s^{-2}
d 1.31×10^{-23} m s^{-2}
8 D **9** **a** D **b** B **c** C **d** A **e** D
10 The direction of the arrowhead indicates the *direction* of the gravitational force and the space between the arrows indicates the *magnitude* of the field. The field lines always point towards the sources of the field.
11 9.76 N kg^{-1} **12** **a** 9.79 N kg^{-1} **b** 100.61%
13 **a** 11.1 N kg^{-1} **b** C **14** 16
15 **a** 3×10^{7} J **b** 4×10^{7} J
c 2000 m s^{-1} or 2 km s^{-1} **d** 3.5 N kg^{-1}
16 9 N kg^{-1} **17** D **18** C **19** 3.5×10^{9} J
20 No. Air resistance will play a major part as the satellite re-enters the Earth's atmosphere.

Chapter 2 Electric and magnetic fields

2.1 Electric fields

WE 2.1.1 5.62×10^{-4} N C^{-1} **WE 2.1.2** 2.16×10^{-18} J on the field

2.1 review

1 C **2** B
3 **a** True. **b** False. **c** False. **d** True. **e** True.
f False. **g** False.
4 1.25×10^{-2} N **5** 1.39 mC **6** 5.72×10^{11} m s^{-2}
7 1200 V **8** 20 electrons
9 **a** work done by the field **b** no work is done
c work done on the field **d** no work is done
e work done on the field **f** work done by the field
10 **a** 1.09×10^{-19} J **b** Work was done on the field.

2.2 Coulomb's law

WE 2.2.1 6.32×10^{-4} N repulsion
WE 2.2.2 $q_1 = +6.35 \times 10^{-10}$ C $q_2 = -6.35 \times 10^{-10}$ C
WE 2.2.3 8.0×10^{5} N C^{-1} away from the charge Q or to the right

2.2 review

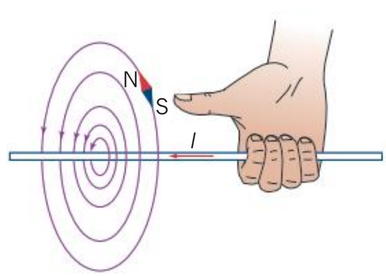
1 **a** positive, repulsion **b** negative, attraction
c negative, repulsion **d** positive, attraction
2 D **3** -8.22×10^{-8} N **4** 1.1×10^{7} N C^{-1}
5 9000 N **6** 1.435 m
7 **a** double, repel **b** quadruple, repel
c double, attract **d** quadruple, repel
8 37 N **9** 1.97×10^{13} electrons

2.3 The magnetic field

WE 2.3.1 The magnetic field direction is perpendicular to the wire. As the current travels along the wire, the magnetic field runs anticlockwise around the wire.

2.3 review

1 B **2** A **3** C **4** C
5



6 A
7 east–away from the north pole of the left hand magnet
8 west–away from the north pole of the right-hand magnet
9 zero
10 **a** A = east, B = south, C = west, D = north
b A = west, B = north, C = east D = south

2.4 Forces on charged objects due to magnetic fields

WE 2.4.1 4.8×10^{-22} N
WE 2.4.2 out of the screen (towards you)
WE 2.4.3 2.5×10^{-3} N per metre of power line
WE 2.4.4 vertically downwards
WE 2.4.5 **a** 0 N **b** 1.0×10^{-3} N out from Santa's house.

2.4 review

1 D
2 **a** South (S) **b** C **c** remains constant
d A **e** Particles with no charge, e.g. neutrons
3 D **4** 0 N **5** no force will apply
6 0.4 N upwards **7** 4.8×10^{-24} N south **8** $2F$ north
9 2.0×10^{-4} N north
10 **a** 0.18 N downwards **b** Same as (a).

2.5 Comparing fields—a summary

2.5 review

1 C **2** B **3** 8.3×10^{-18} m
4 **a** monopoles **b** monopoles and dipoles **c** dipoles
5 radial, static, non-uniform **6** negative **7** D
8 resultant, individual **9** 8.2×10^{-8} N
10 3.6×10^{-47} N

Chapter review

1 0.0225 N **2** D
3 The electrical potential is the work done per unit charge to move a charge from infinity (where the potential is zero) to a point in the electric field. The potential difference, V, is the change in the electrical potential between these two points.
4 25 V **5** C **6** field, charged particle
7 4.17×10^{-18} J **8** 2×10^{-14} N
9 **a** quarter, repel **b** quadruple, repel **c** halve, attract
10 5.42×10^{4} m s^{-1} **11** 0.045 N **12** 45.8 m
13 $+1.63 \times 10^{-4}$ C **14** B into the page
15 $3B$ into the page **16** zero **17** D
18 **a** palm **b** fingers **c** thumb
19 equal to, into **20** 2.78 A
21 **a** 5.0×10^{-9} N into the page **b** 2.0×10^{-3} N into the page
22 9.6×10^{-15} N **23** downwards
24 east–west, as it runs perpendicular to the Earth's magnetic field
25 C **26** 7.9×10^{-6} N **27** 1.9×10^{-48} N

Chapter 3 Applications of fields

3.1 Satellite motion

WE 3.1.1 **a** 589 N **b** 774 N **c** 1150 N
WE 3.1.2 3.08×10^{3} m s^{-1}
WE 3.1.3 **a** 6.70×10^{5} km **b** 1.90×10^{27} kg
c 8.20 km s^{-1}

3.1 review

1 63.7 N **2** 150 N **3** 532 N **4** 441 N
5 B **6** C **7** B **8** D
9 **a** 0.22 m s^{-2} **b** 506 N (or 510 N to 2 significant figures)
10 15.6 days

3.2 DC motors

WE 3.2.1 clockwise as viewed from side 3

3.2 review

1 A **2** 1.0×10^{-2} N into the page
3 1.0×10^{-2} N out of the page **4** 0 N
5 anticlockwise **6** D **7** 2.0×10^{-4} N m
8 0.1 N
9 Current flows into brush P and around the coil from V to X to Y to W. So force on side VX is down, force on side YW is up, so rotation is anticlockwise.
10 D

3.3 Particle accelerators

WE 3.3.1 2.1×10^{7} m s^{-1}
WE 3.3.2 **a** 1.3×10^{5} V m^{-1} **b** 9.4×10^{7} m s^{-1} **c** 1.8×10^{-3} m

3.3 review

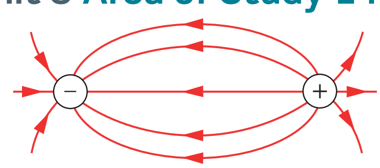
1 B **2** 2.4×10^{-24} N south. **3** 3.0×10^{7} m s^{-1}
4 **a** 9.6×10^{-15} N **b** 4.6×10^{-3} m **5** 9.4×10^{-4} T
6 1.6×10^{6} m s^{-1}
7 A charged particle in a magnetic field will experience a force ($F = qvB$).
As force ∝ velocity, the force will increase as the velocity increases. This will continue while the charge remains in the magnetic field, continuously accelerating the charge.
8 5.8×10^{-2} m or 5.8 cm

Chapter review

1 299 N **2** D **3** D **4** A **5** $11.2T$
6 **a** 0.0540 m s^{-2} **b** 4.38×10^{3} m s^{-1} **c** 5.89 days
7 **a** 0.315 N kg^{-1} **b** 344 m s^{-1}
8 **a** down the page **b** up the page **9** anticlockwise
10 **a** down the page **b** up the page
c zero, because the force is parallel to the coil, rather than perpendicular to it
11 C
12 to reverse the current direction in the coil every half turn to keep the coil rotating in the same direction
13 Electrons are released from a negative terminal (cathode) of the evacuated tube and accelerate under a high potential difference towards a positively charged anode. They hit a fluorescent screen at the rear of the tube.
14 5.9×10^{7} m s^{-1} **15** 2.2×10^{-4} m
16 **a** The electron will curve in an upwards arc from its starting position.
b its velocity and the magnitude of the magnetic field that is acting
17 **a** 1.4×10^{4} V m^{-1} **b** 9.3×10^{6} m s^{-1}
18 4.0×10^{7} m s^{-1} **19** 5.8×10^{-4} T
20 **a** 9.3×10^{-15} N **b** 4.0×10^{-3} m

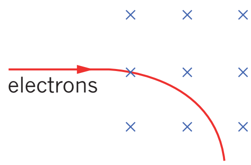
Unit 3 Area of Study 1 review

1



2 2.0 N attraction
3 6.9×10^{9} N C^{-1} to the left (away from the charge)
4 1.05×10^{4} V m^{-1} **5** 1.68×10^{-15} N
6 6.4×10^{-17} J **7** 1.2×10^{7} m s^{-1} **8** 36 V
9 1.5×10^{5} N C^{-1} (or V m^{-1}) downwards **10** 1.28×10^{-18} C
11 8 electrons **12** A **13** 2×10^{-14} N
14 9.9×10^{7} m s^{-1} **15** 1.4×10^{5} V m^{-1}
16 **a** A **b** B **c** G
17 **a** to the left **b** more strongly to the left **c** to the right
18 1×10^{-3} N **19** from west to east **20** 4.9×10^{4} A
21 2×10^{-3} N down **22** B **23** 0.05 N
24 to the right **25** 0.01 N **26** to the left
27 **a** no force **b** no force
28 **a** 0.5 N out of the page **b** 0.5 N into the page
29 The coil will rotate through 90° until the plane of the loop is perpendicular to the field (and the page). It may swing back and forth until it settles in this position.
30 A, B and C

31 **a** The field force on side AB is upwards and that on side CD is downwards.
b When the coil is horizontal.
c When the coil is vertical its momentum will carry it past the true vertical position and then the commutator reverses the direction of the current through the coil and so the forces reverses.
32 **a** 4.0 A **b** $F = 20$ N **c** 64 N **33** 2×10^{-5} m
34



35 5.8×10^{-2} m **36** 2.5×10^{-7} N **37** C
38 D **39** A **40** B **41** 4.25×10^{10} J
42 3.3×10^{3} m s^{-1} **43** **a** 4.0×10^{4} N **b** 8.1×10^{4} N
44 The acceleration increases from 8.1 m s^{-2} to 9.2 m s^{-2}.
45 1.06×10^{5} N **46** 5.5×10^{3} s
47 The mass of the satellite has no effect on its orbital period.
48 60 N **49** 3.6×10^{7} J
50 Determine the energy associated with each grid square by multiplying each area by the mass of 20 kg. Calculate the altitude at which the total area starting from zero height is equal to 40 MJ.

Chapter 4 Electromagnetic induction and transmission of electricity

4.1 Inducing an emf in a magnetic field

WE 4.1.1 8.0×10^{-5} Wb
WE 4.1.2 No, $\varepsilon = 0.695$ V
This is a very small emf and would not be dangerous.
WE 4.1.3 0.50 m s^{-1}

4.1 review

1 A **2** 0 Wb **3** 3.2×10^{-6} Wb
4 The magnetic flux decreases from 3.2×10^{-6} Wb to 0 after one-quarter of a turn. Then it increases again to 3.2×10^{-6} Wb through the opposite side of the loop after half a turn. Then it decreases to 0 again after three-quarters of a turn. After a full turn it is back to 3.2×10^{-6} Wb again.
5 1.3×10^{-5} Wb **6** 0.0144 V or 1.44×10^{-2} V
7 0.84 m s^{-1} **8** 0.10 m **9** 0 V **10** 0.50 V

4.2 Induced emf from a changing magnetic flux

WE 4.2.1 **a** 5.0×10^{-4} Wb **b** 5.0×10^{-3} V
WE 4.2.2 1000 turns

4.2 review

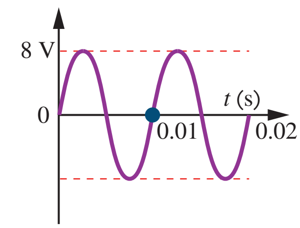
1 1.2×10^{-6} Wb **2** zero **3** 3.0×10^{-5} V
4 C **5** 4×10^{-3} V **6** 2 V **7** 6.0×10^{-3} V
8 The student must induce an emf of 1.0 V in the wire by changing the magnetic flux through the coil. To do this she could change the area of the coil. Required rate of change of flux to produce 1.0 V = 0.01 Wb s^{-1} This would correspond to a change in area from 0.01 m^2 to 0.02 m^2 in a time of 0.1 s. See ProductLink for fully worked solution.
9 0.010 m^2 **10** 0.125 s

4.3 Lenz's law and its applications

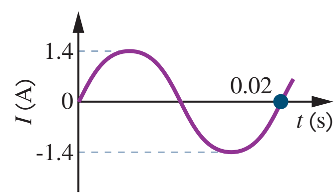
WE 4.3.1 clockwise when viewed from above
WE 4.3.2
(i) through the solenoid from Y to X (through the meter from X to Y)
(ii) no induced emf or current
(iii) through the solenoid from X to Y (through the meter from Y to X)
WE 4.3.3 anticlockwise **WE 4.3.4** 2000 W

4.3 review

1 C **2** **a** A **b** A
3 **a** Anticlockwise
b Any combination of:
1) strength of the magnet
2) speed of the magnet
3) area/diameter of the ring
4) orientation of the ring
5) type of copper making up the ring
6) resistance of the circuit containing the coil.
4 B **5** B
6 $V_p = 8.0$ V
$V_{p\text{-}p} = 16$ V
$V_{rms} = 5.66$ V
7



8



9 29.98 W or 30 W **10** 3.54 A

4.4 Supplying electricity—transformers and large-scale power distribution

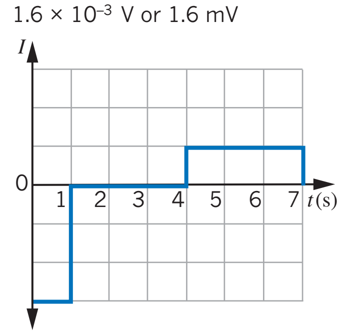
WE 4.4.1 4000 turns **WE 4.4.2** 0.0125 A
WE 4.4.3 3 W **WE 4.4.4** 3.6×10^{5} W or 0.36 MW
WE 4.4.5 500.6 kV

4.4 review

1 B **2** D **3** 40 turns
4 **a** $P_1 = P_2$
b $\frac{I_2}{I_1} = \frac{N_1}{N_2}$
5 **a** 80 V
b 16 W
c 0.20 A
6 **a** 40 turns **b** 0.14 A **c** 24 W **d** A
7 400 W **8** 4×10^{7} W or 40 MW
9 **a** 5000 A **b** 90 kV **10** B

Chapter review

1 **a** 3.2×10^{-3} V or 3.2 mV **b** Clockwise
2 **a** 0.04 V **b** From Y to X.
3 **a** 4×10^{-3} V or 4.0 mV **b** From X to Y.
4 1.6×10^{-3} V or 1.6 mV
5



Either the graph shown or its inversion is correct.
6 1.0 A **7** 10 turns **8** A

9 **a** 18 V **b** 375 W **10** C

11 In a quarter of a turn $\Delta\phi_B = 80 \times 10^{-3} \times 10 \times 10^{-4} = 8 \times 10^{-5}$ Wb

Frequency is 50 Hz so quarter of a turn takes $\frac{1}{4} \times 0.02 = 0.005$ s

$$\varepsilon = -N\frac{\Delta\phi_B}{\Delta t} = 500 \times \frac{8 \times 10^{-5}}{0.005}$$
$$= 8 \text{ V}$$

12 The average emf will double to 16 V.

13 Any two of:
1. Using a DC power supply means that the voltage cannot be stepped up or down with transformers.
2. Hence there will be significant power loss along the 8 Ω power lines.
3. Damage to any appliances operated in the shed that are designed to operate on 240 V AC and not on 240 V DC.

14 clockwise **15** AB and CD **16** 15 A

17 9970 V **18** 450 W **19** See Productlink.

20 anticlockwise

Unit 3 Area of Study 2 review

1 zero

2 Rotate the loop or the magnetic field so they are no longer parallel.

3 When the plane of the loop and the magnetic field direction are perpendicular.

4 0.01 Wb or 10^{-2} Wb

5 As the loop enters the magnetic field there is a flux increasing down through the loop—from Y to X.

6 The loop moves at a speed of 5 cm s^{-1}, and with side length 20 cm, it is halfway into the field, 4×10^{-3} V.

7 8×10^{-3} A **8** 3.2×10^{-5} W

9 The external force that is moving the loop into the magnetic field.

10 Zero, since the loop is totally within the magnetic field and there is no flux change.

11 From X to Y, using the right-hand grip rule, to oppose the decreasing flux down through the loop.

12 5×10^{-6} Wb **13** zero **14** 2.5×10^{-3} V

15 1.25×10^{-3} A

16 No, because there will be no change in flux and therefore no emf generated.

17 200 μA **18** 0.21 T

19 The graph is a sine wave with peak amplitude of 0.9 V and a period of 0.01 s (10 ms).

20 0.64 V

21 The output graph would have half the period and twice the amplitude. The rms voltage would be 1.3 V.

22 50 Hz **23** 200 V **24** B **25** D

26 C **27** 500 Hz **28** 20 V **29** 7.1 V

30 0.71 A **31** 5 W

32 An alternator has a pair of slip rings instead of a split ring commutator.

33 AC is generated in the coils of an alternator. Each slip ring connects to each end of the coil.
The slip rings maintain the AC generated in the coil at the output.

34 0.4 A **35** 6000 V **36** 200 turns

37 849 W **38** 1697 W **39** C **40** C

41 There was a higher current in the power line and hence a voltage drop along the line, leaving a low voltage at the house.

42 218 V, 3488 W

43 At the generator end a 1:20 step-up transformer is required (5000 ÷ 250 = 20). There will be 20 times as many turns in the secondary as in the primary. At the house end a 20:1 step-down transformer is required.

44 0.8 A **45** 1.6 V **46** 1.28 W

47 249.92 V **48** 3998.72 W

49 Losses are about 0.03% of the power generated.

50 The power loss in the power line depends on the square of the current ($P = I^2R$). A higher voltage means a lower current (for the same resistance).

Chapter 5 Newtonian theories of motion

5.1 Newton's laws of motion

WE 5.1.1 a 0 N **b** 6.0 N
c 6.0 N at an angle of 30° below the horizontal.

WE 5.1.2 **a** 1.1×10^4 N in the direction of motion **b** 2.2×10^3 N

WE 5.1.3 **a** $F_\perp = 783$ N, $F_\parallel = 285$ N **b** 783 N
c 3.4 m s^{-2} down the slope

5.1 review

1 No. Phil's inertia made him stay where he was (stationary) as the tram moved forwards. This is an example of Newton's first law. Objects will remain at rest unless a net unbalanced force acts to change the motion.

2 0.098 N upwards **3** **a** 45 N **b** 165 N

4 **a** 1.5 m s^{-2} **b** 120 N **c** 60 N

5 **a** zero **b** 66 N **c** 66 N **6** See Productlink.

7 4.3 N

8 **a** 5.0×10^3 N
b The rope will not break as the tension is less than the breaking strength.

9 **a** A **b** C **c** 490 N up the hill **d** 4.9 m s^{-2}
e Acceleration is not affected by mass if there is no friction.

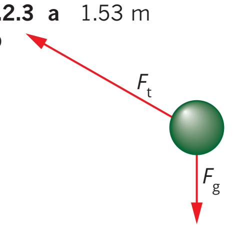
10 A, B and D **11** A

5.2 Circular motion in a horizontal plane

WE 5.2.1 7.5 km h^{-1}

WE 5.2.2 **a** 521 m s^{-2} **b** 3.6×10^3 N

WE 5.2.3 **a** 1.53 m
b



c 2.34 N towards the left **d** 3.05 N

5.2 review

1 B **2** 0.2 s **3** A and D

4 **a** 8.0 m s^{-1} **b** 8.0 m s^{-1} south
c 7.0 m s^{-2} towards the centre, i.e. west

5 8.4×10^3 N west

6 **a** 8.0 m s^{-1} north **b** towards the centre, i.e. east

7 The force needed to give the car a larger centripetal acceleration will eventually exceed the maximum frictional force that could act between the tyres and the road surface. At this time, the car would skid out of its circular path.

8 **a** 2.67 m s^{-2} **b** unbalanced as the skater has an acceleration
c 135 N

9 **a** 28 s **b** 5.0 N

10 **a** 0.5 s **b** 10 m s^{-1} **c** 125 m s^{-2} **d** 310 N

11 **a** 1.2 m
b The forces are her weight acting vertically and the tension in the rope acting along the rope towards the top of the maypole.
c towards point B, the centre of her circular path.
d 170 N towards B **e** 2.6 m s^{-1}

5.3 Circular motion on banked tracks

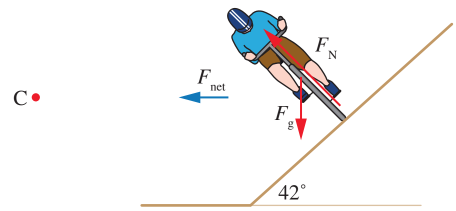
WE 5.3.1 **a** 590 N towards the centre of the circle
b 17 m s^{-1}

5.3 review

1 towards the centre of the circle

2 The architect could make the banking angle larger or increase the radius of the track.

3 Higher up the banked track as the greater speed means that a greater radius is required in the circular path.
4 friction, normal, weight, balanced, normal, weight
5



6 48 km h^{-1}
7 **a** 640 N
b On a horizontal track, F_N is equal and opposite to the weight force, so $F_N = mg = 539$ N. This is less than the normal force on the banked track (643 N).
8 47° **9** **a** 4.9 kN **b** 22°
10 A greater radius will make the car travel higher up the banked track. The driver would have to turn the front wheels slightly towards the bottom of the bank.

5.4 Circular motion in a vertical plane

WE 5.4.1 **a** 4.85 m s^{-1} **b** 15.6 N up
c 3.70 m s^{-1} **d** 6.73 N down

5.4 review

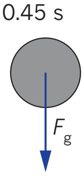
1 **a** It is constant in magnitude. **b** At the bottom of its path.
c At the top of its path. **d** At the bottom of its circular path.
2 3.8 m s^{-1}
3 **a** The weight force from gravity and the normal force from the road.
b 1280 N (or 1.3×10^3 N)
c Yes. When the driver is moving over a hump, the normal force is less than her weight *mg*. Her apparent weight is given by the normal force that is acting and so the driver feels lighter at this point.
d 36 km h^{-1}
4 **a** 31.4 m s^{-1} **b** 19.9 m s^{-1} **c** 8300 N down
5 12.1 m s^{-1} **6** 196 N down **7** 31.3 m s^{-1}
8 188 m s^{-1} **9** **a** 18 m s^{-2} up **b** 1530 N up
10 **a** 9.8 m s^{-2} down. **b** 2.2 m s^{-1}
11 **a** 3.6×10^3 N down **b** 1.3×10^3 N up **c** 7.0 m s^{-1}

5.5 Projectiles launched horizontally

WE 5.5.1 **a** 2.47 s b 49.4 m
c 31.4 m s^{-1} at 50.4° below the horizontal.

5.5 review

1 A and D **2** **a** 1.5 m **b** 7.35 m s^{-1} **c** 7.6 m s^{-1}
3 **a** 0.49 s **b** 2.0 m **c** 9.8 m s^{-2} down
4 **a** 1.0 s **b** 20 m **c** 9.8 m s^{-2} down
d 21.5 m s^{-1} **e** 22.3 m s^{-1}
5 **a** 47 m s^{-1} **b** 58° **6** B and C
7 The hockey ball travels further. A polystyrene ball is much lighter and is therefore more strongly affected by air resistance than the hockey ball.
8 **a** 0.64 s **b** 0.64 s **c** 3.2 m
9 **a** 54 m s^{-1} **b** 22°
10 **a** 10 m s^{-1} forwards **b** 4.4 m s^{-1} down
c 10.9 m s^{-1} at 24° below the horizontal
d 0.45 s **e** 4.5 m
f



5.6 Projectiles launched obliquely

WE 5.6.1 **a** 6.11 m s^{-1} horizontally to the right.
b 0.25 m **c** 0.45 s

5.6 review

1 B **2** 45° **3** 17.3 m s^{-1}
4 **a** 13.6 m s^{-1} **b** 6.34 m s^{-1}
c 9.8 m s^{-2} down **d** 13.6 m s^{-1}
5 **a** 4.0 m s^{-1} **b** 6.9 m s^{-1} **c** 0.70 s
d 3.9 m **e** 4.0 m s^{-1}
6 **a** 24.2 m s^{-1} **b** 24.2 m s^{-1} **c** 24.2 m s^{-1}
7 **a** 14.0 m s^{-1} **b** 4.20 m s^{-1} **c** 5.60 m s^{-1} down
8 24.8 m s^{-1} **9** 28 m s^{-1} **10** 69.2 m **11** C

5.7 Conservation of energy and momentum

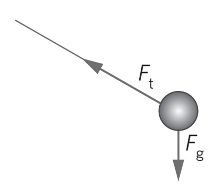
WE 5.7.1 **a** 3.3 m s^{-1} west **b** 2.3×10^4 kg m s^{-1} west
c 2.3×10^4 kg m s^{-1} east
d $\Sigma p_c + \Sigma p_b = -2.3 \times 10^4 + 2.3 \times 10^4 = 0$ Therefore the momentum of the system is constant, as expected.
WE 5.7.2 49.9 m

5.7 review

1 The billiard balls form an isolated system. Momentum is conserved, so if the momentum is zero after the collision, it was initially zero as well. This is possible because the two balls were initially travelling in opposite directions and their momentum vectors cancelled out to give zero.
2 0.41 m s^{-1} east **3** 6.0 m s^{-1}
4 **a** 1.0×10^4 kg m s^{-1} east **b** 1.0×10^4 kg m s^{-1} west **c** 0
5 **a** 0
b It hasn't gone anywhere. The vehicles had a total of zero momentum before the collision and so there still is zero momentum after the collision.
c 1.0×10^4 kg m s^{-1} west
d The change in momentum of the station wagon is
$\Delta p_w = p_{w\,(final)} - p_{w\,(initial)} = 0 - 1.0 \times 10^4$ kg m s^{-1} west
$= 1.0 \times 10^4$ kg m s^{-1} east
6 12 m s^{-1} to the right
7 $v = 5.0$ m s^{-1} in the opposite direction to the shell
8 22 m s^{-1} to the right **9** 2740 J **10** 392 J
11 **a** 33.0 J **b** 33.0 J **c** 21.4 m

Chapter 5 review

1 B **2** **a** 6.5 m s^{-2} **b** 32.5 N
3 4.5×10^3 N **4** **a** 4.9 m s^{-2} **b** $F_N = 0.87\, F_g$
5 **a** 4.9 m s^{-2} **b** 24.5 m s^{-1}
6 **a** 236 N **b** 8.88 m s^{-2} down the ramp
c 506 N down the ramp **d** 9.4 m s^{-1} **e** 506 N up the ramp
7 **a** 508 N **b** 137 N **c** the force of the teenager on the slide
d the force of gravitational attraction from the teenager on the Earth
8 **a** 3.70 m s^{-1} **b** 17.1 m s^{-2} towards the centre of the circle
c 0.430 N
9 **a**



b 0.49 N
10 **a** 2.5 m s^{-2} towards the centre of the circle
b friction between the tyres and the ground
11 **a** 1.02×10^3 m s^{-1} **b** 1.99×10^{20} N
12 3.40×10^{-2} m s^{-2}
13 **a** 10 m s^{-1} south **b** 10 m s^{-1} **c** 13 s **d** 5.0 m s^{-2} west
e 7.5×10^3 N west
14 0.146 m **15** A **16** 15.7 m s^{-1}
17 **a** **(i)** 365 N up **(ii)** 615 N up **b** D
18 Forces acting are: gravity (weight) and the normal force from the base of the bucket on the water.
Both act downwards. Acceleration is towards the centre of the circle, i.e. downwards, and is greater than the acceleration due to gravity.

19 **a** 2.5 m **b** 9.8 m s^{-2} downwards (due to gravity)
20 **a** 10 m s^{-1} **b** 4.4 m s^{-1} **c** 11 m s^{-1}
21 **a** 10.3 m s^{-1} **b** 12.3 m s^{-1} **c** 8.9 m
22 **a** 1.8 J **b** 1.96 J **c** 8.7 m s^{-1}
23 **a** 0 **b** 200 kg m s^{-1} east **c** 200 kg m s^{-1} west
24 1.0 m s^{-1} west **25** 1.68 m s^{-1}
26 **a** 136 kg m s^{-1} west **b** −136 west or 136 kg m s^{-1} east

Chapter 6 Special relativity

6.1 Einstein's theory of special relativity

6.1 review

1 D **2** A and D
3 A hanging pendulum in the spaceship will move from its normal vertical position when the spaceship accelerates.
4 The speed of the ball is greater for Jana than it is for Tom. The speed of the sound is greater forwards than it is backwards for Jana, while for Tom it is the same forwards and backwards. The speed of light is the same for Jana and Tom.
5 **a** 370 m s^{-1} **b** 300 m s^{-1} **c** 360 m s^{-1} **d** 340 m s^{-1}
6 A
7 **a** 10 – 25 = 15 m s^{-1} backwards **b** 3 m backwards **c** 0.2 s
8 **a** 0.1 s **b** 50 m s^{-1} in all frames **c** 1 m
d 50 m s^{-1} as always **e** 0.08 s (approx.)
9 Atomic clocks are precise enough to measure the tiny relativistic differences in time needed to support Einstein's special theory of relativity.
10 very short, μs, very similar, should not, do

6.2 Time dilation

WE 6.2.1 520 s

6.2 review

1 light, oscillation, time, constant
2 'Proper time' is the time measured at rest with respect to the event. Proper times are always less than any other times.
3 1.29 s **4** 48.15 s **5** 2.20 s **6** 1.15 s
7 **a** Simply the height of the clock, 1 m **b** 3.33×10^{-9} s
c ct_c **d** 7.6×10^{-9} s **e** 2.3
8 **a** 1.74×10^{-5} s or 17.4 μs
b Non-relativistic: 655 m
Relativistic: 5178 m
9 $t_0 = 2.93 \times 10^{-11}$ s
10 The equator clock is moving faster relative to the poles. It is also accelerating and hence will run slower. The effect is well below what we can detect as the speed of the equator is 'only' about 460 m s^{-1}, which is about 1.5 millionths of *c*.

6.3 Length contraction

WE 6.3.1 3.90 m **WE 6.3.2** 20.5 m

6.3 review

1 The length that a stationary observer measures in their own frame of reference. That is, the object (or distance) that is being measured is at rest with respect to the observer.
2 A **3** 0.812 m **4** 3.37 m
5 **a** 0.9*c* or 2.71×10^{8} m s^{-1}.
b The fast-moving garage appears even shorter (0.643 m) than its proper length to the car driver.
6 Proper time, t_0, because the observer can hold a stopwatch in one location and start it when the front of the carriage is in line with the watch and stop it when the back of the carriage is in line with it.
7 C
8 **a** 0.866*c* or 2.60×10^{8} m s^{-1} **b** Dan appears half his thickness.
9 23.5 m (At this speed, there is no difference in length.)
10 **a** 1.20 m **b** the proper length, 2.75 m

Chapter 6 review

1 No object can travel at or beyond the speed of light, so the value of $\frac{v^2}{c^2}$ will always be less than 1.
The number under the square root sign will also, therefore, be a positive number less than one.
The square root of a positive number less than one will always be less than one as well.
2 0.000000014 or 1.4×10^{-8}
3 A (postulate 2) and C (postulate 1) **4** At the poles.
5 C
6 Space and time are interdependent—motion in space reduces motion in time.
7 3×10^{8} m s^{-1} **8** A and B **9** B
10 You could not tell the difference between (i) and (iii), but in (ii) you could see whether an object like a pendulum hangs straight down.
11 In your frame of reference time proceeds normally. Your heart rate would appear normal. As Mars is moving at a high speed relative to you, people on Mars appear to be in slow motion as time for them, as seen by you, will be dilated.
12 26.8 s **13** **a** 0.992 s **b** 0.992 s **14** C
15 **a** 0.866*c* or 2.598×10^{8} m s^{-1}
b No, it can't have doubled to over *c*!
$v = 0.968c$ or 2.90×10^{8} m s^{-1}
16 **a** 1.67 s **b** Length: 1.80 m Height: 1.0 m
17 **a** 5.6 years **b** 2.45 years
c Raqu sees the distance as only 2.183 ly
18 **a** 1.4 mm
b No, as the motion is perpendicular to the north–south direction.
19 **a** $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1}{\sqrt{1-\frac{(0.995c)^2}{c^2}}} = \frac{1}{\sqrt{1-(0.995)^2}} = 10.01$
b No, they don't experience any difference in their own time frame.
c About 25.1 years from our frame of reference.
d 2.51 years
e No! They see the distance between Earth and Vega foreshortened because of the high relative speed, so to them the distance is only about 2.5 ly.
20 In the frame of reference of the observer, the muon's time slows so it lives much longer and therefore makes it to the Earth. In the muon's frame of reference, the distance to Earth is contracted so it has a much shorter distance to travel.

Chapter 7 The relationship between force, energy and mass

7.1 Impulse

WE 7.1.1 impulse = 3.94×10^{4} kg m s^{-1} south-west
$F_{ave} = 3.15 \times 10^{3}$ N south-west
WE 7.1.2 101 N s **WE 7.1.3** 5.0×10^{4} N

7.1 review

1 Ball A, ball C, ball B. **2** 38 N **3** 6.89 N s
4 A, C and D **5** 2.14×10^{3} N
6 **a** 1.39×10^{6} kg m s^{-1} **b** 1.39×10^{6} kg m s^{-1}
7 Jacinta is correct. The higher pressure ball would have a greater change in momentum over a shorter time so would exert a greater force.
8 200 N s **9** 1.0×10^{3} N **10** 12.9 N

7.2 Work done

WE 7.2.1 **a** 106 J **b** 46 J **c** 60.0 J **WE 7.2.2** 90 J

7.2 review

1 A **2** 36 J **3** 24 J **4** 12 J
5 0.27 J **6** D **7** 1.8×10^{3} J
8 The magnetic force does no work on the particle.
9 374 J **10** 510 J

7.3 Strain potential energy

WE 7.3.1 **a** 2.0×10^4 N m^{-1} **b** 2.25 J **c** 14.5 J

7.3 review

1 C, B, A
2 Stiff spring constant = 200 N m^{-1}
Weak spring constant = 50 N m^{-1}
3 3.0 J **4** 0.08 m or 8.0 cm **5** 30 N
6 3.75 J **7** 7 J **8** 7 J
9 No. Hooke's law is not obeyed as the force vs distance graph is not a straight line (not linear).
OR
Yes. Hooke's law is obeyed up to a stretch of 0.05 m (i.e. a distance of 0.15 m) on the graph where the line changes from being linear.
10 at the point where the distance = 0.15 m and the force is 30 N

7.4 Kinetic and potential energy

WE 7.4.1 The kinetic energy after the collision is the same as the kinetic energy before the collision. The collision is perfectly elastic. (See Productlink for the fullly worked calculations.)
WE 7.4.2 4.5×10^8 J
WE 7.4.3 **a** 1.4×10^{10} J **b** 6.2 km s^{-1}

7.4 review

1 2.51×10^6 J **2** A and E **3** 1.6×10^7 J
4 5.3×10^8 J **5** 1.5×10^3 m s^{-1} **6** 2.5×10^{11} J
7 A, C and D **8** B and D **9** 6.4 m s^{-1}

7.5 Einstein's mass–energy relationship

WE 7.5.1 **a** 1.56×10^{-21} kg m s^{-1} **b** 0.998*c*
WE 7.5.2 **a** 4 **b** 4.8×10^{-12} J **c** 5.3×10^{-29} kg

7.5 review

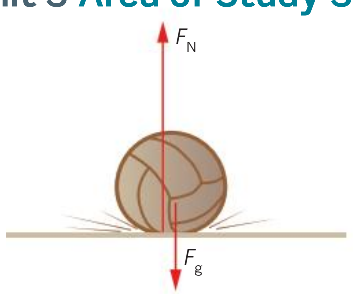
1 9.53×10^5 kg m s^{-1} **2** 9.65×10^{-18} kg m s^{-1}
3 1.59×10^{-23} kg m s^{-1} **4** 1.67×10^{15} J
5 3.11×10^{14} J **6** B **7** 3.59×10^{19} J
8 3.11×10^{31} J **9** C **10** C

Chapter 7 review

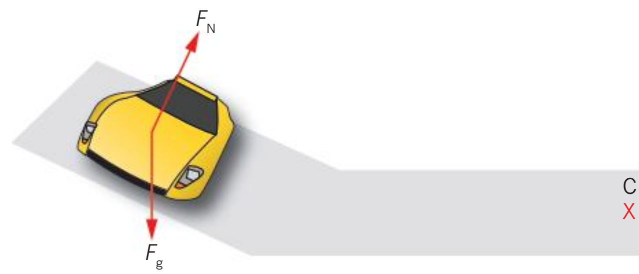
1 B and D **2** 5.3 kg m s^{-1} **3** D
4 **a** Yes, momentum is conserved in all collisions.
b Inelastic; 20 J of kinetic energy has been transformed into heat and sound energy.
c 2 m s^{-1}
5 480 N **6** **a** 3.6×10^4 J **b** 12 m s^{-1}
7 **a** 6.9×10^{10} J **b** 2.6×10^{11} J **8** 3×10^3 J
9 3.20×10^{-2} J **10** A **11** 0.25 m s^{-1}
12 1.10×10^{-2} J **13** 1.1×10^{-2} J
14 **a** more than **b** is not **c** is **d** is
e is not, kinetic energy
15 1.7×10^3 N m^{-1} **16** 5.1×10^{13} J **17** B
18 2.60×10^8 m s^{-1} **19** 3.34×10^{-27} kg
20 1.23×10^{21} J

Unit 3 Area of Study 3 review

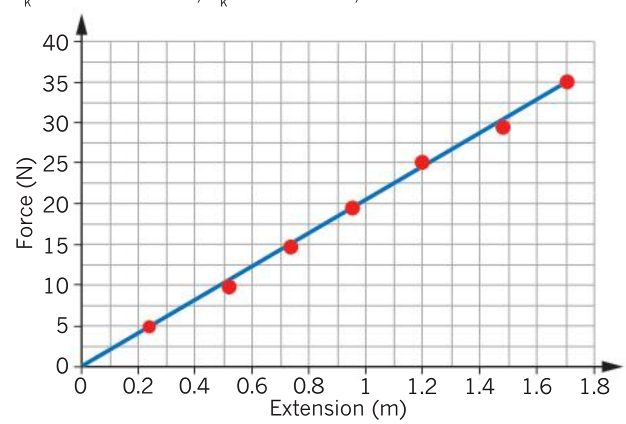
1



2 C **3** unbalanced, balanced **4** 12.6 m s^{-1}
5 31.6 m s^{-2} **6** 1.9×10^3 N **7** 0.6 Hz
8



9 47° **10** 18 m s^{-2} up **11** 1.5×10^3 N
12 1.5×10^3, almost three times larger than the weight force
13 4.9 m **14** 9.8 m s^{-2} down **15** 10.6 m s^{-1}
16 C
17 A. Graph A shows an inelastic collision ($E_{k\ after} < E_{k\ before}$) in which some kinetic energy is transformed into potential energy (the dip) and then back into kinetic energy.
18 D. Momentum is conserved (i.e. is constant) in all collisions.
19 D **20** D **21** 9.8 m s^{-2}
22 2.2 m s^{-1} **23** 0 **24** 5.0 m s^{-1} **25** 1400 J
26 1400 J **27** 3.7 m s^{-1} **28** 200 W
29 24.3 J **30** 4.4 m s^{-1}
31 192 kg m s^{-1} **32** 192 kg m s^{-1}
33 E_k before = 2160 J, E_k after 1597 J, the collision is inelastic.
34



35 20 N m^{-1} **36** 2.25×10^3 J **37** C
38 8.7 m s^{-1} **39** A (postulate 2) and C (postulate 1)
40 A and C **41** C **42** C
43 Aristotle's ideas agreed with our everyday observations. In a space station we would often experience objects moving with constant velocity with no external force, as objects 'floated' around the ship.
44 In your frame of reference time proceeds normally. As Mars is moving at a high speed relative to you, people on Mars appear to be in slow motion as time for them, as seen by you, will be dilated.
45 5.6 years **46** 2.4 years
47 Relative to her, the distance appeared to be foreshortened by the factor γ, thus the distance she travelled was much less than 5 light years.
48 4.23×10^{-12} J **49** 9.3×10^{37} every second
50 3.8×10^{14} kg

Chapter 8 Properties of mechanical waves

8.1 review

1 The particles oscillate back and forth or up and down around a central or average position and pass on the energy carried by the wave. They do not move along with the wave.
2 **a** False: Longitudinal waves occur when particles of the medium vibrate in the *same* direction or *parallel* to the direction of the wave.
b True. **c** True. **d** True.

3 downwards
4 sound, ripples on a pond, vibrations in a rope
5 A has moved right and B has moved left. As the sound wave moves to the right, particles ahead of the compression must move to the left initially to meet the compression and then move forward to carry the compression to the right. Therefore, particle B has moved to the left of its undisturbed position and particle A has now moved to the right of its undisturbed position.
6 C and D
7 The motion of the particles is at right angles (perpendicular) to the direction of travel of the wave itself
8 Longitudinal: a and d Transverse: b and c
9 Mechanical waves move energy via the interaction of particles. The molecules in a solid are closer together than those in a gas. A smaller movement is needed to transfer energy and, hence, the energy of the wave is usually transferred more quickly in a solid when compared with other states of matter.
10 Towards the right, from the tuning fork towards X

8.2 Measuring mechanical waves

WE 8.2.1 0.4 m
WE 8.2.2 amplitude = 0.1 m, period = 0.5 s, frequency = 2 Hz
WE 8.2.3 7.5×10^{14} Hz **WE 8.2.4** 1.3×10^{-15} s

8.2 review

1 **a** C and F **b** wavelength **c** B and D **d** amplitude
2 Wavelength λ = 1.6 m, amplitude = 20 cm
3 **a** 0.4 s **b** 2.5 Hz **4** 6.5 m s^{-1}
5 **a** True.
b False: The period of a wave is *proportional* to its wavelength.
c True.
d False: The wavelength *and frequency* of a wave determine its speed.
6 **a** wavelength = 4 cm; amplitude = 0.5 cm
b 2 cm s^{-1} or 0.02 m s^{-1} **c** red
7 5×10^{-6} s
8 As the speed of each vehicle is the same and there is no relative motion of the medium, the frequency observed would be the same as that at the source.
9 The apparent frequency increases when travelling towards you and decreases when travelling away from you.

8.3 Wave interactions

8.3 review

1 The wave is reflected and there is a 180° change in phase.
2 amplitude
3 **a** True.
b False: As the pulses pass through each other, the interaction *does not* permanently alter the characteristics of each pulse.
c True.
4 B
5 An object subjected to forces varying with its natural oscillating frequency will oscillate with increasing amplitude. This could continue until the structure can no longer withstand the internal forces and fails.
6 52° **7** B
8 Normal walking results in a frequency of 1 Hz or 1 cycle per second i.e. two steps per second. This frequency may result in an increase in the amplitude of oscillation of the bridge over time, which could damage the structure.
9 C

8.4 Standing waves in strings

WE 8.4.1 **a** 0.25 m **b** 0.17 m

8.4 review

1 No, it is a common misconception that standing waves somehow remain stationary. It is only the pattern made by the amplitude along the rope that stays still at the nodes. The rope is still moving, especially at the antinodes.
2 A transverse wave moving along a slinky spring is reflected from a fixed end. The reflected and original waves superimpose, and the resulting pattern of constructive and destructive interference creates a standing wave.
3 0.8 m **4** 1.5 m
5 $\frac{1}{4}$ of the fundamental wavelength **6** 2.5 m
7 0.74 m **8** 300 Hz **9** 600 Hz **10** 900 Hz

Chapter 8 review

1 The particles on the surface of the water move up and down as the waves radiate outwards carrying energy away from the point on the surface of the water where the stone entered the water.
2 Similarities: both are waves, both carry energy away from the source, both are caused by vibrations.
Differences: transverse waves involve particle displacement at right angles to the direction of travel of the wave; longitudinal waves involve particle displacement parallel to the direction of travel of the wave.
3 U is moving down and V is momentarily stationary (and will then move downwards).
4 0.300 m s^{-1} **5** 0.044 m **6** 5 m
7 1.1 m **8** 256 Hz **9** C and D
10 The frequency would increase and the velocity would remain unchanged.
11 It undergoes a phase change.
12 **a** transmission **b** reflection **c** absorption
13 The green wave
14 Sound waves are longitudinal mechanical waves where the particles only move back and forth around an equilibrium position, parallel to the direction of travel of the wave. When these particles move in the direction of the wave, they collide with adjacent particles and transfer energy to the particles in front of them, losing kinetic energy in the process.
15 If an object is made to vibrate at its resonant frequency, the amplitude of its vibrations will increase with time. If a building or bridge is made to resonate by, e.g., wind matching its natural frequency, the resulting vibrations could damage the structure.
16 100 Hz **17** 300 Hz **18** 0.50 Hz
19 0.091 m or 9.1 cm **20** All of the options are correct.
21 When there is relative motion between the source, observer and medium

Chapter 9 The nature of light

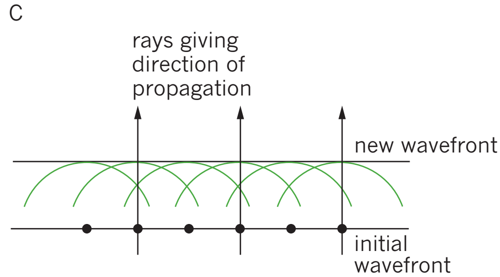
9.1 Light as a wave

WE 9.1.1

WE 9.1.2 1.52 **WE 9.1.3** 1.62×10^{8} m s^{-1}
WE 9.1.4 28.2° **WE 9.1.5** 24.4°

9.1 review

1 **a** wave model **b** wave model **c** particle model
2 C
3



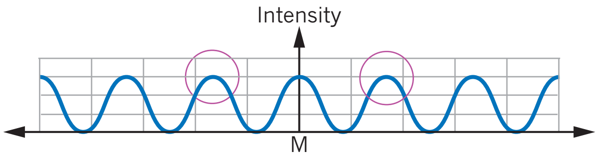
4 slower than 5 2.17×10^8 m s^{-1} 6 1.31
7 35.3° 8 b and c 9 D
10 Polarisation is a phenomenon in which transverse waves are restricted in their direction of vibration. Polarisation can only occur in transverse waves and cannot occur in longitudinal waves. Since light can be polarised, it must be a transverse wave.

9.2 Interference: Further evidence for the wave behaviour of light

WE 9.2.1 550 nm

9.2 review

1 D 2 C and D 3 A and D
4



5 Up until Young's experiment, most scientists supported a particle or 'corpuscular' model of light. Young's experiment demonstrated interference patterns, which are characteristic of waves. This led to scientists abandoning the particle theory and supporting a wave model of light.
6 **a** increase **b** decrease **c** increase
7 2610 nm or 2.61×10^{-6} m
8 **a** destructive **b** constructive **c** destructive
9 1400 nm 10 455 nm

Section 9.3 Electromagnetic waves

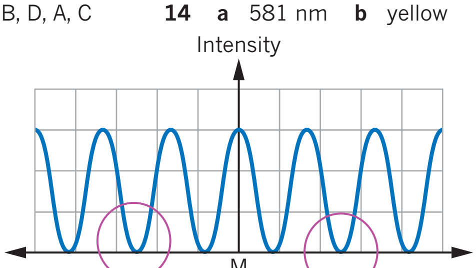
WE 9.3.1 5.0×10^{14} Hz

9.3 review

1 B 2 D 3 D
4 X-rays, visible light, infrared radiation, FM radio waves
5 **a** 4.57×10^{14} Hz **b** 5.09×10^{14} Hz
c 6.17×10^{14} Hz **d** 7.56×10^{14} Hz
6 0.07% 7 500 nm 8 4.3 m
9 1.5×10^{18} Hz 10 0.122 m

Chapter 9 review

1 A
2 The diffraction pattern would spread out more from blue to green.
3 D
4 Both snow and water reflect light. This reflected light is known as glare. The light reflected from water and snow is partially polarised. Both snowboarders and sailors are likely to wear polarising sunglasses as these will absorb the polarised glare from the snow or water respectively.
5 2.25×10^8 m s^{-1}
6 increases, away from
7 **A:** incident ray **B:** normal **C:** reflected ray
D: boundary between media **E:** refracted ray
8 2.1×10^8 m s^{-1} 9 28.9°
10 **a** 32.0° **b** 53.7° **c** 21.7° **d** 1.97×10^8 m s^{-1}
11 **a** 19.5° **b** 19.1° **c** 0.4° **d** 1.96×10^8 m s^{-1}
12 **a** 49.8° **b** 40.5° **c** 27.6°
13 B, D, A, C 14 **a** 581 nm **b** yellow
15



16 radio waves, microwaves, infrared, visible, ultraviolet, X-rays, gamma rays
17 **a** microwaves **b** infrared waves **c** X-rays 18 490 m
19 Young shone monochromatic light on a pair of narrow slits. He identified that the resulting pattern of bright and dark fringes corresponded to regions of constructive and destructive interference, which could only be explained by a wave model of light.
20 A microwave oven produces electromagnetic waves with the same frequency as the resonant frequency of water molecules. This makes the water molecules in food vibrate, and the energy is transferred to the rest of the food, heating it up.

Chapter 10 Light and matter

10.1 The photoelectric effect and the dual nature of light

WE 10.1.1 2.4×10^{-19} J **WE 10.1.2** 1.5 eV
WE 10.1.3 1.5 eV **WE 10.1.4** 5.0 eV **WE 10.1.5** 2.07 eV

10.1 review

1 **a** 3.03×10^{-19} J **b** 3.38×10^{-19} J
c 4.09×10^{-19} J **d** 5.01×10^{-19} J
2 If light shining on it causes electrons to be released.
3 **a** True.
b False: When light sources of the same intensity but different frequencies are used, the higher frequency light has a higher stopping voltage, but it produces the same maximum current as the lower frequency.
c True.
4 **a** 4.1 eV **b** 4.6 eV **c** 6.2 eV 5 D
6 0.066 eV 7 0.25 eV 8 C and D
9 **a** True.
b False: The stopping voltage is reached when the photocurrent is reduced completely to zero.
c True. **d** True.
10 1.68 eV

10.2 The quantum nature of light and matter

WE 10.2.1 5.7×10^{-13} m **WE 10.2.2** 1.0×10^{-36} m
WE 10.2.3 0.17 nm **WE 10.2.4** 1.47×10^{-27} kg m s^{-1}

10.2 review

1 7.3×10^{-10} m 2 1.8×10^5 m s^{-1} 3 B
4 **a** 3.5×10^{-11} m **b** 2.1×10^7 m s^{-1}
5 The wavelength of a cricket ball is so small that its wave-like behaviour could not be seen by a cricket player.
6 Because the radius of an atom is smaller than the wavelength of an electron.
7 1.32×10^5 m s^{-1}
8 $W = qV = \frac{1}{2}mv^2$

$$v = \sqrt{\frac{2qV}{m}}$$

$$\lambda = \frac{h}{mv}$$

$$= \frac{h}{m\frac{\sqrt{2qV}}{\sqrt{m}}}$$

$$= \frac{h}{\sqrt{2qVm}}$$

9 $\lambda = \frac{h}{mv}$

$$\lambda mv = h$$

$$mv = \frac{h}{\lambda}$$

$$p = \frac{h}{\lambda}$$

10 Because a high-speed electron has a shorter wavelength than a light wave.

10.3 Light and matter

WE 10.3.1 4.6×10^{-19} J **WE 10.3.2** Lyman

WE 10.3.3

A photon of 6.7 eV corresponds to the energy required to promote an electron from the ground state to the second excited state ($n = 1$ to $n = 3$). The photon may be absorbed.

A photon of 5.0 eV cannot be absorbed.

A photon of 11.0 eV may ionise the mercury atom. The ejected electron will leave the atom with 0.6 eV of kinetic energy.

10.3 review

1 The electrons in a sample become excited when the substance is heated or an electric current flows through it. As the electrons return to their ground state, a photon is emitted.

2 4.0×10^{-19} J **3** 3.0×10^{-6} m

4 **a** Light Emitting Diode
b Light Amplification by Stimulated Emission of Radiation

5 675 nm **6** 12.75 eV

7 9.74×10^{-8} m or 97.4 nm

8 de Broglie proposed a model where electrons were viewed as matter waves with wavelengths that formed standing waves within an atomic orbit circumference. A bowed violin string forms standing waves between the bridge of the violin and the violinist's finger.

9 High-energy orbits of multi-electron atoms, the continuous emission spectrum of solids and the two close spectral lines in hydrogen that are revealed at high resolution.

10 −0.54 eV

10.4 Heisenberg's uncertainty principle

10.4 review

1 It would *increase*.

2 The location of electrons can't be restricted to specific orbital paths. While electrons exist at particular energy levels, it is impossible to know the precise location of the electron.

3 The predictions of Newtonian physics do not fit with Heisenberg's uncertainty principle and hence the wave–particle duality observed at the sub-atomic scale.

4 The photon would transfer energy to the electron, changing its momentum and hence changing its path.

5 For the normal-sized world around us, the inclusion of Planck's constant, h, in the measure of uncertainty means that the level of uncertainty in determining the position of everyday objects is extremely small.

6 There are no constraints on the path of the light. The uncertainty in the position of a photon becomes large and hence the uncertainty in momentum becomes small.

7 The uncertainty in the position of the electron is increased. As a consequence, the uncertainty of the momentum of the electron will decrease. Fringes on the diffraction pattern will move closer together.

8 While an uncertainty can always be calculated, when applied to large objects the uncertainty is insignificant. The uncertainty principle is applied to sub-atomic particles in the study of quantum mechanics.

9 D

Chapter 10 review

1 2.5 eV **2** 8.0×10^{-19} J **3** photoelectrons

4 1.2×10^{15} Hz **5** 2.9 eV **6** 1.95 eV

7 Rb = 2.1 eV, Sr = 2.5 eV, Mg = 3.4 eV, W = 4.5 eV

8 **a** See Productlink for the plotted graph.
b 4.1×10^{-15} eV s **c** 5.0×10^{14} Hz
d No. The frequency of red light is below the threshold frequency for rubidium.

9 **a** 4.78 keV
b The electrons have a de Broglie wavelength which is similar to the wavelength of the X-rays. This is evidence for the dual nature of light and matter.
c 2.6×10^{-24} kg m s^{-1}

10 **a** a sequence of maximum and minimum intensities
b The electrons are exhibiting wave-like behaviour. Electrons are not light but, like light, a beam of electrons can be diffracted.

11 B **12** 1.7×10^{-35} m **13** No

14 Energy levels in an atom cannot assume a continuous range of values but are restricted to certain discrete values, i.e. the levels are quantised.

15 2.9×10^{15} Hz

16 Bohr's work on the hydrogen atom convinced many scientists that a particle model was needed to explain the way light behaves in certain situations.

17 The colours that are missing in the absorption spectrum match the colours that are visible in the emission spectrum.

18 As the filament heats up, the free electrons in the tungsten atoms collide, accelerate and emit photons.

19 It will increase.

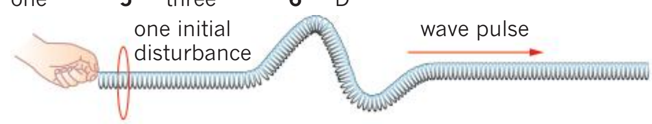
20 The uncertainty in the electron's position would increase.

Unit 4 Area of Study 1 review

1 C **2** $A_x : A_y = 2:1$ **3** $f_x : f_y = 1:2$

4 one **5** three **6** D

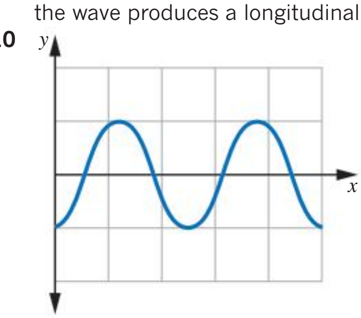
7



8 A mechanical wave involves energy being transferred from one location to another, without any net transfer of matter.

9 Displacement of the medium perpendicular to the direction of travel of the wave produces a transverse wave.
Displacement of the medium parallel to the direction of travel of the wave produces a longitudinal wave.

10



11 4.0 m **12** A **13** C

14 1.3 m **15** 2 cm

16 The wave is reflected with a decrease in amplitude and a phase change. Energy is transformed into heat by the fixed end.

17 The resulting waveform will be the vector addition of the individual waves due to the principle of superposition. The shape, amplitude or speed of the individual waves is not altered.

18 Amplitude—twice the original; wavelength—same as the original.

19 No, they will have the same characteristics as before the interaction.

20 110 Hz **21** 330 Hz **22** 550 Hz

23 A = node, B = antinode

24 They show the maximum and minimum positions of the rope as it oscillates.

25 Three; the third.

26 The apparent frequency will increase when the source is moving

27 Being pushed in the direction of motion once, at the correct moment each oscillation, will drive the swing at its resonant frequency and will result in a gain in amplitude.

28 3×10^{8} m s^{-1} **29** 3×10^{-8} m **30** ultraviolet

31 For example:
- UV lamps are used to sterilise surgical equipment in hospitals
- UV lamps are used to sterilise food and drugs
- UV rays help the body to produce vitamin D.

32 infrared **33** C and D **34** 11.9°

35 1.24×10^{8} m s^{-1}

36 They have different wavelengths (or frequencies).

37 They would need to pass white light through a triangular glass prism.

38 The white light will separate into the component colours which, from top to bottom, will be red, orange, yellow, green, blue, indigo, violet.
39 The electromagnetic variations occur in only one direction at right angles to the direction of propagation.
40 Sound is a longitudinal wave.
41 When the polarising axes are parallel.
42 Maximums when the polarising axes are parallel: 0° and 180°; zeros when the polarising axes are perpendicular: 90° and 270°.
43 Δx will be doubled. **44** Δx will be doubled.
45 Δx will be doubled.
46 There will be a wider central band.
47 The wave model and the particle (or corpuscular) model.
48 Young's experiment resulted in bright and dark bands or fringes being seen on a screen. These can only be due to interference effects and a wave model.
49 A and B. At A, two crests add to give maximum constructive interference. At B, two troughs add to give maximum constructive interference.
50 500 nm

Unit 4 Area of Study 2 review

1 B **2** C **3** A **4** 5 eV
5
- Only certain frequencies of light will emit photoelectrons.
- There is no time difference between the emission of photoelectrons by light of different intensities.
- The maximum kinetic energy of the ejected photoelectrons is the same for different light intensities of the same frequency.

6 2.42×10^{-8} m **7** A series of bright and dark fringes.
8 They are exhibiting wave-like behaviour. **9** 164 m s^{-1}
10 $f = 2.13 \times 10^{15}$ Hz $\lambda = 1.41 \times 10^{-7}$ m
11 $f = 1.18 \times 10^{15}$ Hz $\lambda = 2.53 \times 10^{-7}$ m
12 $f = 5.07 \times 10^{14}$ Hz $\lambda = 5.91 \times 10^{-7}$ m
13 1.04×10^{-17} J **14** 4.78×10^{6} m s^{-1}
15 1.52×10^{-10} m
16 If incident light had an energy value less than the minimum energy difference between the lowest and next orbital levels within the hydrogen atom, the light would not result in any orbital changes. Therefore the light would not be absorbed by the atom.
17 0.0416 nm
18 There would be circular bands or fringes of specific spacing around a common central point.
19 Greater momentum –> lower wavelength (using $\lambda = \frac{h}{p}$), so the circular bands would be more closely spaced.
20 de Broglie would say that the electrons were diffracted as they passed through the gaps between the atoms in the crystal, creating a diffraction pattern.
21 Electrons have a de Broglie wavelength. Their orbit must fit an integral number of wavelengths so that a standing wave is formed ($2\pi r = n\lambda$). Only energy levels corresponding to these wavelengths exist.
22 B, C, D, E
23 The energy that an electron gains when moved through a potential of 1 V.
24 Any excess energy results in extra kinetic energy of the electron.
25 The discovery that light can display both particle and wave properties was mirrored when electrons were found to have wave properties, when moving very fast, as well as particle properties.
26 C
27 Each incident photon interacts with only one electron; therefore, the energy of the emitted electrons will depend only on the frequency of the incident light (according to $E - hf$) and not on the intensity (the number of incident photons).
28 Altering the intensity of the light corresponds to waves of greater amplitude, which should deliver more energy to the electrons and, therefore, the emerging electrons should have higher energy. (This is not observed.)
29 Photon energy > ionisation energy
30 6.4×10^{-20} J **31** 3.41×10^{-25} kg m s^{-1}
32 1.94×10^{-9} m
33 Since there is no energy level 10.0 eV above the ground state, the photon cannot be absorbed.
34 1.8 eV, 4.9 eV and 6.7 eV **35** 1.85×10^{-7} M
36 3.04×10^{-8} m
37 $E_{k\,max}$ is the maximum kinetic energy of the emitted electrons. f is the frequency of the light incident on the metal plate. ϕ is the work function i.e. the minimum energy required to eject an electron.
38 $E_{k\,max}$ is not altered. **39** More current will flow.
40 3 **41** 2.21×10^{-22} kg m s^{-1}
42 They must have equivalent wavelengths.
43 3.6×10^{-11} m **44** 3.6×10^{-11} m
45 1.8×10^{-23} kg m s^{-1}
46 No. The energy of the X-rays is given by $E = \frac{hc}{\lambda}$ and the energy of the electrons is given by $\Delta E_k = \frac{1}{2} mv^2$.
47
1. It predicts a minimum frequency (threshold frequency and energy) before electrons are emitted. (The wave model predicts that any frequency should work.)
2. The energy of the emitted electrons depends only on the frequency of the incident light. (The wave model predicts that increasing the intensity of light would increase the energy of the emitted electrons.)
3. It also explains an absence of any time delay before electrons are emitted when weak light sources are used. (This time delay is suggested by the wave model.)

48 **a** B **b** A **c** C **d** D
49 Narrowing the slit makes the diffraction pattern become wider, as Heisenberg's uncertainty principle predicts (less uncertainty in position means more uncertainty in momentum).
50 The uncertainty in a particle's position, Δx, becomes greater, as the right-hand side of the relation is constant.

Glossary

A

absorb To take in (energy).

absorption spectrum Spectrum containing dark lines in the positions of the wavelengths that are absorbed by a gas as light passes through it. This is related to the emission spectrum of the gas.

acceleration due to gravity Rate at which a falling object will accelerate in a gravitational field. Equivalent to the gravitational field strength. Measured in m s^{-2}.

aether An invisible, massless, rigid substance that was proposed as the medium in which light waves propagate. There is no experimental evidence for the existence of the aether.

air resistance A retarding force that acts in the opposite direction to the motion of an object or projectile.

alternator An electric generator that produces alternating current (AC).

altitude Height above a planet's surface.

amplitude The absolute value of the maximum displacement from a zero value during one period of an oscillation.

angle of incidence The angle that a ray of light or wave, meeting a surface, makes with a normal to the surface at the point of meeting.

angle of reflection The angle that a ray of light or wave, reflected from a surface, makes with a normal to the surface at the point of reflection.

antinode The region of maximum amplitude between two adjacent nodes in a standing wave or interference pattern. See also *node*.

apparent weight The weight felt by a person when their body is stationary or in motion. Sometimes it is higher or lower than their usual weight. Equivalent to the size of the normal reaction force acting on the person.

apparent weightlessness When an object is in free fall and there is no force between it and its surroundings. It appears to be weightless although it is still under the influence of gravity. When there is no normal reaction force acting between an object and a surface.

armature A revolving structure in an electric motor or generator, wound with the coils that carry the current. It rotates within a magnetic field to induce an emf.

artificial satellite Body such as Sputnik, the Hubble Space Telescope, or NOAA-19, made by humans and placed in orbit around a planet or the Moon.

B

banked track A track inclined at some angle to the horizontal enabling vehicles to travel at higher speeds when cornering compared with around a horizontal curved path.

breaking point The point indicating the force and extension at which a material fails on a force vs extension curve.

brushes Devices that transfer the current in the rotating coil to a stationary external circuit by pressing against the split ring commutator or the slip rings.

C

cathode ray tube A vacuum tube in which a hot cathode emits a beam of electrons that pass through a high voltage anode and are focused or deflected before hitting a fluorescent screen.

centripetal acceleration Acceleration directed towards the centre of a circle when an object moves with constant speed in a circular path.

centripetal force The force that causes an object to travel in a circular path; can include gravity, tension, normal force and friction.

classical physics The physics of Galileo and Newton, in which the addition of velocities has no limit, and length and time are constant.

coherent Waves that are in phase i.e. at the same stage at the same time.

collision An interaction in which two or more objects exert forces on one another, causing an exchange of energy between them. It is not an absolute requirement for the objects or particles to physically make contact. Therefore, the interaction of two charged particles, which repel one another without ever physically touching, is also a collision.

compression To press or squeeze, as in the area of increased pressure within a longitudinal wave.

conserved Not created or destroyed, but remaining constant.

constructive interference The process in which two or more waves of the same frequency combine to reinforce each other. The amplitude of the resulting wave is equal to the sum of the amplitudes of the superimposed waves.

crest The highest part or top of a transverse wave.

critical angle The angle of incidence that produces an angle of refraction of 90°. The largest angle for which refraction will occur; at angles larger than the critical angle, light undergoes total internal reflection.

D

de Broglie wavelength Wavelength associated with a particle due to its motion.

deformation Change in shape of an object as a result of the application of a force. This is often used to describe a permanent change, when work is done to change the structure of the material.

design speed Relating to a banked track, the speed at which a vehicle experiences no sideways force as it travels around track. It is dependent on angle.

destructive interference The process in which two or more waves of the same frequency combine to cancel each other out. The amplitude of the resulting wave is equal to the difference between the amplitudes of the superimposed waves.

diffraction The bending of waves around obstacles or through gaps in their path.

diffraction pattern The pattern of dark and light bands that is seen when light passes through a single small gap. Areas of constructive interference appear as bright bands and areas of destructive interference appear as dark bands.

diffuse Spread out, scattered widely or thinly.

dipole Two electric charges or magnetic poles that have equal magnitudes but opposite signs, usually separated by a small distance.

direct current A continuous electric current that flows in one direction only, without substantial variation in magnitude. Batteries are a source of direct current. Abbreviated to DC.

dispersion The process of splitting light into its component colours to create a spectrum.

Doppler effect The apparent change in frequency of a wave for an observer due to the relative motion between the observer and the source.

E

elastic Describes a material that returns to its original shape after being deformed.

elastic collision Collision in which kinetic energy is conserved.

elastic limit The limit of force applied and deformation that can occur for which the material will behave elastically. Further force will result in permanent deformation.

electric field A region of space where charged objects experience a force due to the field created by another charged object.

electric field strength A measure of the force per unit charge on a charged object within an electric field, with the units N C^{-1}. Field strength can also be a measure of the difference in electrical potential per unit distance, with the units V m^{-1}.

electrical potential The work required per unit charge to move a charged object from infinity to a point in the electric field, with the units J C^{-1}.

electromagnet A magnet consisting of an iron or steel core wound with a coil of wire, through which a current is passed. The core only becomes magnetised when current is flowing.

electromagnetic induction The creation of an electric current, or an emf, in a loop of wire as the result of changing the magnetic flux through the loop.

electromagnetic radiation Energy emitted in continuous waves with two transverse, mutually perpendicular components: a varying magnetic field and a varying electric field.

electromagnetic spectrum The range of all possible frequencies of electromagnetic radiation. The visible spectrum is just one small part of the electromagnetic spectrum.

electron gun Uses a heated cathode to produce an electron beam and a series of charged plates to accelerate the beam.

electron-volt Amount of energy equal to the charge of an electron multiplied by 1 volt, i.e. $1 \text{ eV} = 1.6 \times 10^{-19} \times 1 = 1.6 \times 10^{-19}$ J. An alternative to the joule as a unit in which to measure energy.

emission spectrum Spectrum of coloured lines in the positions of the wavelengths of light emitted when a gas is heated or has an electric current passed through it. This is related to the absorption spectrum of the gas.

excited state Higher energy state of an atom above the ground state ($n > 1$).

F

Faraday's law Law stating that the average emf generated in a coil is proportional to the rate of change of magnetic flux and the number of turns in the coil.

field A region of space around an object where a force can be felt by other objects.

field lines A two-dimensional graphic representation of a field, using arrows to indicate the direction of the field. The closer the field lines, the stronger the field.

frame of reference A coordinate system that is usually fixed to a physical system that contains an object and/or an observer. There can be a frame of reference within another frame of reference.

free fall A motion whereby gravity is the only force acting on a body.

frequency The number of waves passing a given point in one second or number of repeats of a cycle each second. Measured in hertz (Hz).

fundamental The harmonic with the lowest frequency and simplest form of vibration, which has only one antinode.

G

Gedanken German word for 'thought'. Einstein used this term to describe his theoretical 'experiments' on relativity.

generator An electrical device that converts kinetic energy into direct current (DC) electricity. Usually, a coil is rotated causing it to cut across a magnetic field.

geostationary satellite Satellite that remains in orbit above the same place on the Earth's surface. It has the same period as the Earth's rotation, i.e. 24 hours. Only occurs at an altitude of 36 000 km above the Earth.

gravimeter Sensitive instrument used by geologists to detect small variations in gravitational field strength.

gravitational constant, G Universal constant of value 6.67×10^{-11} N m^2 kg^{-2}.

gravitational field The region around an object where other objects will experience a gravitational force.

gravitational field strength The strength of gravity, usually measured at the surface of a planet. Equivalent to the acceleration due to gravity, g. Measured in newtons per kilogram (N kg^{-1}).

gravitational force The force of attraction acting between two objects that have mass.

gravitational potential energy The energy that a body possesses due to its position in a gravitational field. A scalar quantity that is measured in joules (J).

ground state Lowest energy state of an atom ($n = 1$).

H

harmonic A frequency that is a whole number multiple of the same basic frequency.

Heisenberg's uncertainty principle Concept that any measurement of a system creates a disturbance of the system with a resulting uncertainty in the measurement.

I

ideal transformer Where the input power and the output power are equal and the transformer is 100% efficient. Real transformers obtain close to this value.

impulse The change in the momentum of an object. Can be calculated as the difference between the final and initial momentum. It can also be calculated for collisions by multiplying the average force by the duration of the interaction, or by finding the area under a force–time graph.

incandescent Emission of light due to very high temperature.

inclined plane Sloping surface or ramp.

induced current Electric current produced by changing a magnetic flux in the region of a conductor or by moving the conductor in a magnetic field.

inelastic collision Describes a collision in which kinetic energy is not conserved.

inertial frame of reference A frame of reference that is either moving with a constant velocity or is stationary. It is not accelerating.

interference The process in which two or more waves of the same frequency combine to reinforce or cancel each other out. The amplitude of the resulting wave is equal to the sum of the amplitudes of the superimposed waves. See also *constructive interference* and *destructive interference*.

inverse square law Relationship between two variables where one is proportional to the reciprocal of the square of the other.

isolated system Situation where there should only be internal forces acting between the objects and no interaction with objects outside the system.

K

kinetic energy The energy of a moving body, measured in joules (J). Kinetic energy is a scalar quantity.

L

laser Source of a narrow beam of intense, monochromatic, polarised, coherent radiation.

law of conservation of energy Energy cannot be created or destroyed, but can only be changed or transformed from one form to another.

law of conservation of momentum In any collision or interaction between two or more objects in an isolated system, the total momentum of the system will remain constant. That is, the total initial momentum is equal to the total final momentum.

length contraction Length in a moving frame of reference appears shorter when viewed by a stationary observer.

Lenz's law A law stating that the direction of the induced current in a conductor is such that its associated magnetic field opposes the change in flux that caused it.

light-emitting diode (LED) Semiconductor diode that uses the excitation of electrons to emit light.

longitudinal Extending in the direction of the length rather than across something. The vibrations of a longitudinal wave are in the same direction as, or parallel to, the direction of travel of the wave.

Lorentz factor $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$

Lorentz force The force experienced by a point charge moving along a wire that is in a magnetic field; the force is at right angles to both the current and the magnetic field. Named for the Dutch physicist who shared a 1902 Nobel Prize for researching the influence of magnetism on radiation.

M

magnetic Of or relating to magnetism or magnets. Having the properties of a magnet. Capable of being magnetised or attracted by a magnet.

magnetic field A magnetic field is a region influenced by a magnet or something with the properties of a magnet.

magnetic flux The strength of a field in a given area expressed as the product of the area and the component of the field strength at right angles to the area (i.e. $\Phi = B_{\perp}A$

magnetic flux density Amount of magnetic flux per unit area. In other words, it describes 'the closeness of magnetic field lines'. Same as magnetic field strength.

magnetic pole Magnetic poles are two limited regions in a magnet at which the field of the magnet is most intense.

magnitude The size of a quantity without regard for its direction.

mechanical energy The energy that a body possesses due to its position or motion. Kinetic energy, gravitational energy and elastic potential energy are all forms of mechanical energy.

mechanical wave A mechanical wave is a wave that propagates as an oscillation of matter, and therefore transfers energy through a medium.

medium A physical substance, such as air or water, through which a mechanical wave is propagated.

metal vapour lamp Lamp that contains a low-pressure gas that becomes excited and emits photons with the colour characteristic of the element in the gas, e.g. sodium vapour lamp.

mnemonic A mnemonic device is any learning technique that aids information retention. Mnemonics aim to translate information into a form that the brain can retain better than its original form. Even the process of learning this conversion might aid in the transfer of information to long-term memory.

momentum The product of the mass and the velocity of an object. Momentum is a vector. It is measured in kg m s^{-1}.

monochromatic Light of a single colour, e.g. red light.

monopole A single mass or point electric charge. A mass is considered to be a monopole at its centre of mass. Magnetic poles only exist, as far as we currently know, as dipoles.

N

natural satellite A body such as the Moon or a planet (not made by humans) that is in orbit around another body.

Newton's law of universal gravitation Law that states that the attractive gravitational force between two masses is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.

node A point at which the amplitude of two or more superimposed waves has a zero or minimum value. See also *antinode*.

normal A line constructed at 90° (i.e. perpendicular) to a surface at the point that a wave strikes the surface.

normal reaction force Force with which a surface pushes back on an object, at right angles to the surface. Same as the apparent weight of an object. Symbol F_N or N, measured in newtons.

nuclear fusion Reaction in which nucleons are joined to form a new species of nucleus, with a release of energy.

O

overtone Any of the higher-level harmonics, except for the fundamental frequency.

P

paradox A situation that appears to have contradictory elements.

particle accelerator A machine that can accelerate a charged particle (proton, electron) or an atomic nucleus to very high speeds, including speeds that approach the speed of light.

path difference The difference in the lengths of the paths from each slit to the screen in a double-slit experiment.

period The time interval taken to complete one cycle of a regularly repeating phenomenon, such as a rotating object or in a sound wave. The SI unit for period is seconds (s).

phase The fraction of a cycle of a wave that has been completed at a specific point in time, usually expressed as an angle. A particular stage in a periodic process such as a wave.

photocurrent Current caused by the flow of photoelectrons during the photoelectric effect.

photoelectric effect Spontaneous emission of electrons from a metal surface when it is illuminated by light of particular frequencies and energies.

photoelectron An electron released from an atom due to the photoelectric effect.

photon Packet or bundle of electromagnetic radiation (light).

plane wave A constant-frequency wave whose wavefronts are infinite parallel lines or planes. A straight wave.

point charge An ideal situation in which all of the charge on an object is considered to be concentrated at a single point. The point size is negligible in relation to the distance between it and another point charge.

polarisation The phenomenon in which transverse waves are restricted in their direction of vibration.

pole The north pole of a freely suspended magnet is attracted to the Earth's magnetic North Pole (a magnetic south). The south pole is attracted to the Earth's magnetic South Pole (a magnetic north). See *magnetic pole*.

postulate A suggestion that is put forward as a fact as a basis for further discussion or reasoning.

potential difference The work required per unit charge to move a charged object between two points in the electric field, with the units J C^{-1} or volts (V).

projectile Object moving freely through the air without an engine or power source driving it.

proper length A measurement of length made from the frame of reference in which the object being measured is stationary.

proper time A measurement of time made with a clock that doesn't move relative to the point at which the start and end of the event occurs.

pulse A short section of a wave that is not continuous or repetitive.

Q

quantum Plural quanta. According to the quantum model, electromagnetic radiation is emitted from objects as discrete packets called quanta. Each quantum has an energy proportional to its frequency according to the equation $E = hf$.

quantum mechanics Area of modern particle physics where the wave properties of electrons are studied.

R

rarefaction An area of decreased pressure within a longitudinal wave.

ray The straight line path of a wave drawn perpendicular to a wave front. (Also a narrow beam of light.)

reflect To cause light, other electromagnetic radiation, sound, particles or waves, to bounce back after reaching a boundary or surface.

refraction The bending of light, sound, or other type of wave, in passing at an angle to the surface from one medium into another in which its wave speed is different.

refractive index An index or number that is allocated to a medium indicating its refracting properties; ratio of the speed of light in a vacuum, c, to the speed of light in the medium, v, i.e. $n = \frac{c}{v}$.

resonance The state of a system in which an abnormally large vibration is produced in response to an external vibration. Resonance occurs when the frequency of the vibration is the same, or nearly the same, as the natural vibration frequency of the system.

right-hand rules 1. The right-hand grip rule tells us the direction of the magnetic field (curled fingers) around a current (thumb). 2. The right-hand force rule tells us the force (palm) on a current (thumb) in a magnetic field (straight fingers).

root mean square The square root of the arithmetic mean of the squares of the numbers in a given set of numbers. In terms of alternating power, the root mean square value, $P_{rms} = \frac{P_{peak}}{2}$. Similarly, for voltage and current. Alternatively, it is the effective mean (average) value of an AC supply.

S

satellite Object in a stable orbit around a central body. Could be natural, like a planet, or artificial, like a communications satellite.

simultaneous When two events occur at exactly the same time.

sinusoidal Having a magnitude that varies as a sine curve.

slip rings Components of alternators (AC generators) that allow a constant electrical connection to be made between the rotating armature and the static external circuit through which the generated alternating current flows.

Snell's law Describes the relationship between incident light and refracted light in two media. The ratio of the refractive indices of the two materials is equal to the ratio of the sines of the angle of incidence and the angle of refraction.

solar wind A continuous stream of charged particles ejected by the Sun. It consists mostly of protons and electrons and has enough energy to escape the Sun's gravitational field, at speeds ranging from about 300 to 800 km s^{-1}, which allows it to reach the Earth in about 3.9 days. The speed and intensity of the charged particles depend on magnetic activity at different regions of the Sun.

solenoid A coil of wire that acts as an electromagnet when electric current is passed through it due to the magnetic field that is set up by the current passing through it. Solenoids are often used to control the motion of metal objects, such as the switch of a relay.

spacetime A term used to describe the situation in which the 3-dimensional space coordinate system is linked to the 1-dimensional time system.

split ring commutator a component of DC generators and motors that typically resembles a ring that has been cut into two equal pieces or shells. Each part of the ring has a fixed connection to the ends of the coil, while also making contact with the stationary brushes. This means the connection between the rotating coil and the static circuit is reversed every half turn, which ensures the direction of current in the circuit is constant (in the case of the generator) or the direction of rotation is constant (in the case of the motor).

standing wave Also called a stationary wave, the periodic disturbance in a medium resulting from the combination of two waves of equal frequency and intensity travelling in opposite directions.

stator A portion of a machine that remains stationary with respect to rotating parts, especially the collection of stationary parts in the magnetic circuits of a motor or generator.

step-down transformer Device that decreases the secondary voltage compared to the primary voltage.

step-up transformer Device that increases the secondary voltage compared to the primary voltage.

stopping voltage The applied voltage required to stop all photoelectrons from reaching the collector electrode. For a particular frequency of incident light on a particular metal, the stopping voltage is a constant.

strain potential energy The energy stored in a material when it is stretched or compressed. If the material is elastic, this energy can be returned to the system, but in inelastic materials permanent change occurs.

superposition When two or more waves travel in a medium, the resulting wave at any moment is the sum of the displacements associated with the individual waves.

synchrotron Large particle accelerator in a circular shape producing a very intense, very narrow beam of electromagnetic radiation called synchrotron light.

T

tangential Describes a direction forming a tangent to a curve.

threshold frequency The minimum frequency of electromagnetic radiation for which the photoelectric effect can occur for a given material.

time dilation When one observer watches events in a frame of reference that is moving (very fast) relative to him/her, time in that frame of reference will appear to go more slowly. People in the moving frame do not experience any difference in the rate at which time passes. This effect is one of the strange consequences of Einstein's theory of special relativity.

torque Any force or system of forces that causes or tends to cause rotation. A turning or twisting force effect.

torsion balance Device used to measure very small twisting forces. Cavendish used this device to measure the force of attraction between lead balls held a small distance apart.

total internal reflection Occurs when the angle of incidence exceeds the critical angle for refraction. Light or waves are reflected back into the medium; there is no transmission of light.

transform To change form, as in energy changing form from stored energy in a spring to kinetic energy as the spring is released.

transformer A device that transfers an alternating current from one circuit to one or more other circuits, usually with an increase (step-up transformer) or decrease (step-down transformer) in voltage. The input goes to a primary coil; the output is taken from a secondary coil or windings linked by induction to the primary coil.

transmit To cause light, heat, or sound, etc. to pass through into a medium.

transverse Lying or extending across something. The vibrations of a transverse wave are at right angles to the direction of travel of the wave.

trough The lowest part or bottom of a transverse wave.

U

uniform Constant, unvarying.

V

voltaic pile An early form of battery consisting of a pile of paired plates of dissimilar metals, such as zinc and copper, each pair being separated from the next by a pad moistened with an electrolyte (mild acid). Also called galvanic pile or Volta's pile.

W

wave front The set of points reached by a wave or vibration at the same instant. Wave fronts generally form a continuous line or surface.

wavelength The distance, measured in the direction of travel of a wave, between two successive points at the same phase.

wave–particle duality The theory that, in some experiments, light and matter behave like waves and, in other experiments, they behave like particles.

weight Force due to gravity that acts on any mass in a gravitational field. Symbol F_g or W, measured in newtons.

work Transfer or transformation of energy. Work is done when a force causes a displacement in the direction of the force.

work function The energy required to remove an electron from its state of being bound to an atom; measured in joules or electron-volts.

Index

Note: Page numbers in **bold** refer to the main discussion of key terms in the text.

Acknowledgements

We thank the following for their contributions to our text book. The following abbreviations are used in this list: t = top, b = bottom, l = left, r = right, c = centre.

COVER: Science Photo Library, David Parker

123RF: 1xpert, p. 15tr; claudiodivizia, p. 312c; donatas1205, p. 357t; Alexandra Guseva, p. 244; Patricia Hofmeester, p. 220t; jewhyte, p. 156tr; Denys Kuvaiev, p. 240t; lightpoet, p. 249; Felix Lipov, p. 202t; Rafael Angel Irusta Machin, p. 157b; manganganath, p. 208; Daniel Prudek, p. 20; Roman Sigaev, p. 356lb; tristan3d, p. 85.

AAP: Adam Davy/PA, p. 167tr; Yukihiro Ugeda, p. 228.

Age Fotostock: David Nunuk, p. 348.

Alamy Stock Photo: Action Plus Sports Images, p. 231l; Agencja Fotograficzna Caro, p. 233b; Alchemy, pp. 47tr, 47tl; Stephen Barnes/Sport, p. 179b; Oliver Burston, p. 364t; Adam Burton, pp. 271, 299, 369-374; ClassicStock, p. 257tl; dpa Picture-Alliance, pp. 189tr, 227, 264; European Pressphoto Agency B.V. (EPA), p. 189tl; Garry Gay, pp. 71, 103-6; Geo-grafika, p. ix; Paul Hebditch, p. 167br; H.S. Photos , pp. 346c, 346t; imageBROKER, p. 118l; Interfoto, p. 361; Magictorch, pp. 201, 265-9; mauritius images GmbH, p. 287l; Photos 12, p. 339t; Prisma Archivo, p. 82bl; sciencephotos, p. 284lt; Marmaduke St. John, p. 51; Sueddeutsche Zeitung Photo, p. 340b; VintageCorner, p. 2r; Dave Wilson, pp. 147, 200; Zuma Press, Inc, p. 161t.

Australian Synchrotron: pp. 93t, 99.

Brookhaven National Laboratory: Courtesy Brookhaven National Laboratory, p. 93b.

Burrows, Keith: p. 45t.

Cross, Malcolm: pp. 309bc, 309rt.

DK Images: p. 309bl; Stephen Oliver, p. 47bl.

EcoSwitch®: p. 134.

European Space Agency: pp. 19, 82t; ESA/AOES Medialab, p. 79b.

Fotolia: p. 65b, BillionPhotos.com, p. 356lt; Joggie Botma, p. 32c; ChrisVanLennepPhoto, p. 274br; Firma V, p. 295; Michael Ireland, p. 32b; itestro, p. 276t; Craig Jewell, p. 32t; Mr Twister, p. 287rt; oceandigital, p. 325b; petejau, pp. 31, 70; rueangwit, p. 288; Science Photo, p. 272.

Getty Images: p. 54l; Andrew Brookes, National Physical Laboratory, p. 212t; Dr Tony Brain/Science Photo Library, pp. 343bl, 343bc, 343br; Steve Christo, p. 231c; Tim Clayton, p. 183; Paul Ellis, p. 231r; Emilio Segre Visual Archives, p. 349b; Julian Finney, p. 230lt; Bellingham Herald/Contributor, p. 15tl; Gary Hincks/Science Photo Library, p. 5r; moodboard, p. 78l; David Parker/IMI/Univ. of Birmingham High TC Consortium, p. 124; Photo Researchers, p. 96c; Quinn Rooney, p. 165r; Royal Astronomical Society, p. 11r; Royal Institution of Great Britain, p. 108l; Science Photo Library, pp. 94t, 114, 117bl, 148, 179t, 209; Clive Streeter, p. 3; Stringer, p. 138; Sheila Terry, p. 5l; Trevor Clifford Photography/Science Photo Library, pp. 48bt, 57; Detlev Van Ravenswaay/Science Photo Library, p. 56t; William West/AFP, p. 168l.

ImageFolk (was Amana/Corbis): Duomo, p. 235; Heritage Images, p. 89t; p. 330t; Tetra Images, p. 300t.

Leung, Terry: p. 310b.

NASA Images: courtesy of nasaimages.org: pp. 11l, 17b, 21, 72r, 257tr; Graphic: Dixon Rohr, p. 46r; ESA/Hubble Telescope, p. 80; ISS Expedition 12 Crew, p. 81l; JPL, p. 10.

Newspix: Craig Greenhill, p. 172.

Retrospect Photography: Dale Mann, pp. 275lc, 275lt, 275lb, 33l (t-b), 379t, 379c, 379b, 380, 384, 386t, 386b.

Science Photo Library: pp. 63, 202b, 223l, 324b; Andrew Lambert Photography, pp. 327l, 308b; Martin Bond, p. 54r; Carlos Clarivan, p. 350; Henning Dalhoff, p. 344; DESY, p. 358l; Monique Feist/Paul Tafforeau/ESRF/Pascal Goetgheluck, p. 358r; Mark Garlick, p. 79t; Russell Kightley, p. 315b; LMSAL/Stanford University/NASA, p. 325t; Walt Myers, p. 253; N.A. Sharp, NOAO/NSO/KITT PEAK FTS/AURA/NSF, p. 346b; David Parker, p. 339b; Sputnik, p. 77t; Trevor Clifford Photography, p. 302t; Mark Williamson, pp. 270, 314t.

Science Source - Photo Researchers, Inc: GIPhotoStock, pp. 317r, 318r.

Shutterstock: Miss Kanithar Aiumla-Or, p. 132lt; aniad, p. 252; ArtisticPhoto, pp. 107, 143-6; Gunnar Assmy, p. 186; Arvind Balaraman, p. 34r; Dedyukhin Dmitry, p. 233t; Elnur, p. 323b; Filipe Frazao, p. 4; FreshPaint, p. 191; Arina P Habich, p. 176l; Iryna1, p. 2l; Brian Kinney, p. 170r; saied shahin kiya, p. 324t; koya979, p. 353r; langdu, p. 309rb; Magdanatka, p. 329; MilanB, p. 45br; Mopic, p. 307l; MrSegui, p. 282; muzsy, p. 156bl; PeterPhoto123, p. 306r; Pichitchai, p. 132lb; pistolseven, p. 290; Keith Publicover, p. 6; Plati Photography, p. 377; pukach, p. 40; Bruce Rolff, pp. x, 1, 30; Atiketta Sangasaeng, p. 307r; Carlos E. Santa Maria, p. 347; Sideways Design, p. 69r; solarseven, p. 320t; Thieury, p. 232t; Tristan3D, p. 17t; Taras Vyshnya, p. 41b; Westend61 Premium, p. 168c; zstock, p. 336.

The Picture Source: Terry Oakley, pp. 342tla, 342tlc, 342tlb.

Twostaricebox: p. 123b.

Victorian Curriculum and Assessment Authority (VCAA): Selected extracts from the VCE Physics Study Design (2017-2021) are copyright Victorian Curriculum and Assessment Authority (VCAA), reproduced by permission. VCE® is a registered trademark of the VCAA. The VCAA does not endorse this product and makes no warranties regarding the correctness or accuracy of its content. To the extent permitted by law, the VCAA excludes all liability for any loss or damage suffered or incurred as a result of accessing, using or relying on the content. Current VCE Study Designs and related content can be accessed directly at www.vcaa.vic.edu.au pp. ix, 2, 31, 71, 107, 147, 201, 227, 270-1, 299, 329.