PHYSICS 1

PRELIMINARY COURSE

THIRD EDITION

PHYSICS 1

PRELIMINARY COURSE

THIRD EDITION

MICHAEL ANDRIESSEN • GRAEME LOFTS
RIC MORANTE • JOSEPH BARRY MOTT
CONSULTANT: YOKA McCALLUM

jacaranda plus

JACARANDA HSC SCIENCE

Third edition published 2009 by
John Wiley & Sons Australia, Ltd
42 McDougall Street, Milton, Qld 4064

First edition published 2002
Second edition published 2004

Typeset in 10.5/12 pt New Baskerville

National Library of Australia
Cataloguing-in-publication data

Title:	Physics 1: preliminary course/ Michael Andriessen [et al.]
Edition:	3rd ed.
ISBN:	978 0 7314 0820 7 (pbk.)
Series:	Jacaranda HSC science
Notes:	Includes index.
Target audience:	For secondary school age.
Subjects:	Physics — Textbooks.
Other authors/contributors:	Andriessen, Michael.
Dewey number:	530

Illustrated by the Wiley Art Studio, Craig Jackson, Stephen Francis, Paul Lennon

Printed in China by
Printplus Limited

10 9 8 7 6 5 4

CONTENTS

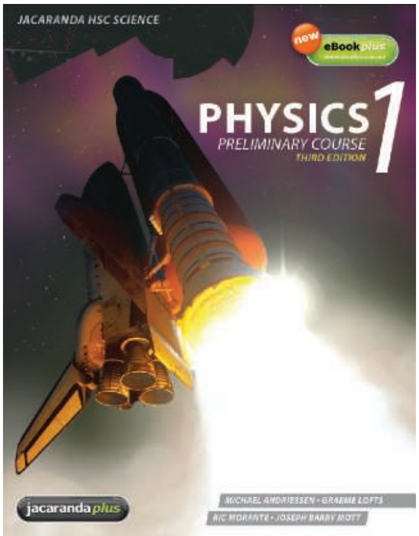
PRELIMINARY MODULE

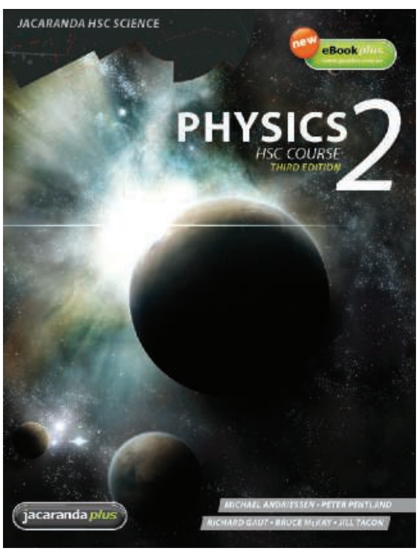
Electrical energy in the home

PRELIMINARY MODULE

Moving about

PREFACE





This third edition of *Physics 1: Preliminary Course* is published in response to amendments made to the Stage 6 Physics syllabus for Year 11 students in New South Wales in 2003 and beyond. This text, which is part of the successful *Jacaranda HSC Science* series, meets all the requirements of the amended content, has been thoroughly updated, and establishes a solid grounding in concepts that are further developed in the revised edition of the Year 12 book, *Physics 2: HSC Course*.

The text covers the four Preliminary core modules in detail in a full-colour presentation, supported by stimulating and contemporary images and clearly labelled diagrams. The content focuses carefully and comprehensively on the syllabus points and a syllabus table is provided on pages xi–xiv to help students and teachers locate the particular pages of the book where content is covered.

A range of features is included that aim to stimulate students' interest and build on their knowledge and understanding:

- *Key terms* are shown in bold coloured type and the definition is given in the margin. A comprehensive glossary of all key terms is included at the end of the book.
- *Sample problems* with worked solutions are interspersed to help students understand and apply the formulae as specified in the syllabus.
- *Physics in focus* and *Physics fact* boxes explain developments and highlight applications of this dynamic science in the contemporary world. *Websites* are included where relevant for extending students' interest.
- *Chapter reviews* consist of:
  - a *Summary* that lists the main points of the content, for revision and to reinforce learning
  - a wide range of descriptive and numerical *Questions* that enable students to address the syllabus outcomes.
- The mandatory *Practical activities* are included at the ends of chapters to meet syllabus needs. Within the chapters, an icon in the margin indicates a relevant point or context where each practical activity might be attempted.
- The *Appendices* include a formulae and data sheet, the periodic table, the list and definitions of key words used in HSC examination questions, and answers to the numerical questions.
- Complete answers and worked solutions for the questions in both *Physics 1: Preliminary Course* and *Physics 2: HSC Course* are available online at the JacarandaPLUS website (www.jacplus.com.au).

The authors and consultant for *Physics 1: Preliminary Course* and *Physics 2: HSC Course* are all experienced teachers and educators. They have approached the topics with a keen sense of responsibility to students to provide clear and accurate explanations and examples of the principles of physics in the context of real-world observations, up-to-date information and cutting-edge technology. They hope that students using this text will meet the challenges and enjoy the rewards of the course while developing a deeper appreciation of the relevance and fascination of physics in their everyday lives.

Next generation teaching and learning

This title features eBookPLUS: an electronic version of the textbook and a complementary set of targeted digital resources. These flexible and engaging ICT activities are available to you online at the JacarandaPLUS website (www.jacplus.com.au).

eBookPLUS icons within the text direct students to the online resources, which include:

- *eModelling:* Excel spreadsheets that provide examples of numerical and algebraic modelling
- *eLessons:* Video and animations that reinforce study by bringing key concepts to life
- *Interactivities:* Interactive study activities that enhance student understanding of key concepts through hands-on experience
- *Weblinks:* HTML links to other useful support material on the internet.

jacaranda plus

Next generation teaching and learning

About eBookPLUS

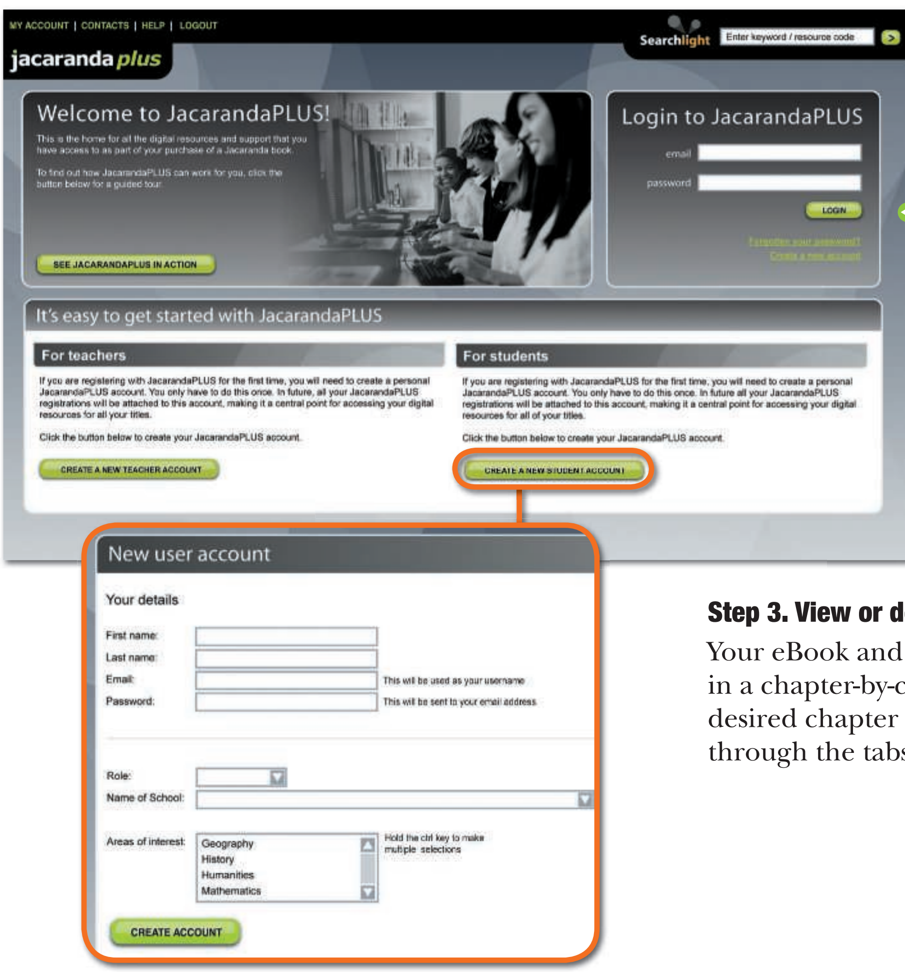
Physics 1: Preliminary Course, 3rd edition features eBookPLUS: an electronic version of the entire textbook and supporting multimedia resources. It is available for you online at the JacarandaPLUS website (**www.jacplus.com.au**).

Using the JacarandaPLUS website

To access your eBookPLUS resources, simply log on to **www.jacplus.com.au**. There are three easy steps for using the JacarandaPLUS system.

Step 1. Create a user account

The first time you use the JacarandaPLUS system, you will need to create a user account. Go to the JacarandaPLUS home page (**www.jacplus.com.au**) and follow the instructions on screen.



Step 2. Enter your registration code

Once you have created a new account and logged in, you will be prompted to enter your unique registration code for this book, which is printed on the inside front cover of your textbook.

LOGIN

Once you have created your account, you can use the same email address and password in the future to register any JacarandaPLUS books.

Step 3. View or download eBookPLUS resources

Your eBook and supporting resources are provided in a chapter-by-chapter format. Simply select the desired chapter from the drop-down list and navigate through the tabs to locate the appropriate resource.

Minimum requirements

- Internet Explorer 7, Mozilla Firefox 1.5 or Safari 1.3
- Adobe Flash Player 9
- Javascript must be enabled (most browsers are enabled by default).

Troubleshooting

- Go to the JacarandaPLUS help page at **www.jacplus.com.au**
- Contact John Wiley & Sons Australia, Ltd.
  Email: support@jacplus.com.au
  Phone: 1800 JAC PLUS (1800 522 7587)

SYLLABUS GRID

Module 1: THE WORLD COMMUNICATES (chapters 1–4, pages 1–73)

1. The wave model can be used to explain how current technologies transfer information

Students learn to:	pages	*Students:*	pages
• describe the energy transformations required in one of the following: – mobile telephone – fax/modem – radio and television	9	• perform a first-hand investigation to observe and gather information about the transmission of waves in: – slinky springs – water waves – ropes or use appropriate computer simulations	5, 7, 8, 13, 14
• describe waves as a transfer of energy disturbance that may occur in one, two or three dimensions, depending on the nature of the wave and the medium	5–6	• present diagrammatic information about transverse and longitudinal waves, direction of particle movement and the direction of propagation	7, 8, 10, 11–12
• identify that mechanical waves require a medium for propagation while electromagnetic waves do not	7	• perform a first-hand investigation to gather information about the frequency and amplitude of waves using an oscilloscope or electronic data-logging equipment	14, 17–19, 26–7
• define and apply the following terms to the wave model: ‘medium’, ‘displacement’, ‘amplitude’, ‘period’, ‘compression’, ‘rarefaction’, ‘crest’, ‘trough’, ‘transverse waves’, ‘longitudinal waves’, ‘frequency’, ‘wavelength’, ‘velocity’	3–4, 7–8	• present and analyse information from displacement–time graphs for transverse wave motion	10, 11, 12
• describe the relationship between particle motion and the direction of energy propagation in transverse and longitudinal waves	7–8	• plan, choose equipment for and perform a first-hand investigation to gather information to identify the relationship between the frequency and wavelength of a sound wave travelling at a constant velocity	12, 24–7
• quantify the relationship between velocity, frequency and wavelength for a wave: $v = f\lambda$	4	• solve problems and analyse information by applying the mathematical model of $v = f\lambda$ to a range of situations	4, 11, 12

2. Features of a wave model can be used to account for the properties of sound

Students learn to:	pages	*Students:*	pages
• identify that sound waves are vibrations or oscillations of particles in a medium	16–17	• perform a first-hand investigation and gather information to analyse sound waves from a variety of sources using the cathode-ray oscilloscope (CRO) or an alternative computer technology	18–19, 23, 26–7
• relate compressions and rarefactions of sound waves to the crests and troughs of transverse waves used to represent them	17–18	• perform a first-hand investigation, gather, process and present information using a CRO or computer to demonstrate the principle of superposition for two waves travelling in the same medium	26–7
• explain qualitatively that pitch is related to frequency and volume to amplitude of sound waves	18–19	• present graphical information, solve problems and analyse information involving superposition of sound waves	21–3, 25, 26–7
• explain an echo as a reflection of a sound wave	19–20		
• describe the principle of superposition and compare the resulting waves to the original waves in sound	20–3		

3. Recent technological developments have allowed greater use of the electromagnetic spectrum

Students learn to:	pages	*Students:*	pages
• describe electromagnetic waves in terms of their speed in space and their lack of requirement of a medium for propagation	29–30	• plan, choose equipment or resources for and perform a first-hand investigation and gather information to model the inverse square law for light intensity and distance from the source	38, 44, 45
• identify the electromagnetic wavebands filtered out by the atmosphere, especially UV, X-rays and gamma rays	33–6	• analyse information to identify the waves involved in the transfer of energy that occurs during the use of one of the following: – mobile phone – television – radar	40, 41
• identify methods for the detection of various wavebands in the electromagnetic spectrum	30–2	• analyse information to identify the electromagnetic spectrum range utilised in modern communication technologies	30–2
• explain that the relationship between the intensity of electromagnetic radiation and distance from a source is an example of the inverse square law: $I \propto \frac{1}{d^2}$	37–8		
• outline how the modulation of amplitude or frequency of visible light, microwaves and/or radio waves can be used to transmit information	38–40		
• discuss problems produced by the limited range of the electromagnetic spectrum available for communication purposes	37, 40		

4. Many communication technologies use applications of reflection and refraction of electromagnetic waves

Students learn to:	pages	*Students:*	pages
• describe and apply the law of reflection and explain the effect of reflection from a plane surface on waves	47	• perform first-hand investigations and gather information to observe the path of light rays and construct diagrams indicating both the direction of travel of the light rays and a wavefront	69–71
• describe ways in which applications of reflection of light, radio waves and microwaves have assisted in information transfer	54–5	• present information using ray diagrams to show the path of waves reflected from: – plane surfaces – concave surfaces – convex surface – the ionosphere	48, 69 51–4, 70–1 51–2, 70–1 55
• describe one application of reflection for each of the following: – plane surfaces – concave surfaces – convex surfaces – radio waves and being reflected by the ionosphere	48–9, 53–5		
• explain that refraction is related to the velocities of a wave in different media and outline how this may result in the bending of a wavefront	56–7	• perform an investigation and gather information to graph the angles of incidence and refraction for light encountering a medium change showing the relationship between these angles	71–3
• define refractive index in terms of changes in the velocity of a wave in passing from one medium to another	57–8	• perform a first-hand investigation and gather information to calculate the refractive index of glass or perspex	71–3
• define Snell’s Law: $\frac{v_1}{v_2} = \frac{\sin i}{\sin r}$	57–9	• solve problems and analyse information using Snell’s Law	58–9, 68–9
• identify the conditions necessary for total internal reflection with reference to the critical angle	60–1		
• outline how total internal reflection is used in optical fibres	61–2		

5. Electromagnetic waves have potential for future communication technologies and data storage technologies

Students learn to: • identify types of communication data that are stored or transmitted in digital form	pages 62–5	*Students:* • identify data sources, gather, process and present information from secondary sources to identify areas of current research and use the available evidence to discuss some of the underlying physical principles used in one application of physics related to waves, such as: – global positioning system – CD technology – the internet (digital process) – DVD technology	pages 62–7

Module 2: ELECTRICAL ENERGY IN THE HOME (chapters 5–8, pages 75–169)

1. Society has become increasingly dependent on electricity over the last 200 years

Students learn to: • discuss how the main sources of domestic energy have changed over time • assess some of the impacts of changes in, and increased access to, sources of energy for a community • discuss some of the ways in which electricity can be provided in remote locations	pages 80–6 78, 88, 90 86	*Students:* • identify data sources, gather, process and analyse secondary information about the differing views of Volta and Galvani about animal and chemical electricity and discuss whether their different views contributed to increased understanding of electricity	pages 77–90

2. One of the main advantages of electricity is that it can be moved with comparative ease from one place to another through electric circuits

Students learn to: • describe the behaviour of electrostatic charges and the properties of the fields associated with them • define the unit of electric charge as the coulomb • define the electric field as a field of force with a field strength equal to the force per unit charge at that point $\boldsymbol{E} = \frac{\boldsymbol{F}}{q}$ • define electric current as the rate at which charge flows (coulombs/second or amperes) under the influence of an electric field • identify that current can be either direct, with the net flow of charge carriers moving in one direction, or alternating, with the charge carriers moving backwards and forwards periodically • describe electric potential difference (voltage) between two points as the change in potential energy per unit charge moving from one point to the other (joules/coulomb or volts) • discuss how potential difference changes at different points around a DC circuit • identify the difference between conductors and insulators • define resistance as the ratio of voltage to current for a particular conductor: $R = \frac{V}{I}$ • describe qualitatively how each of the following affects the movement of electricity through a conductor: – length – cross-sectional area – temperature – material	pages 92–102 93 98–9 104–6 115 108 109–10 94 110–12 112–15, 120–1	*Students:* • present diagrammatic information to describe the electric field strength and direction: – between charged parallel plates – about and between a positive and negative point charge • solve problems and analyse information using $\boldsymbol{E} = \frac{\boldsymbol{F}}{q}$ • plan, choose equipment for and perform a first-hand investigation to gather data and use the available evidence to show the relationship between voltage across and current in a DC circuit • solve problems and analyse information applying $R = \frac{V}{I}$ • plan, choose equipment for and perform a first-hand investigation to gather data and use the available evidence to show the variations in potential difference between different points around a DC circuit • gather and process secondary information to identify materials that are commonly used as conductors to provide household electricity	pages 100–2 98–9 119–21 110–12 119–20 114

3. Series and parallel circuits serve different purposes in households

Students learn to: • identify the difference between series and parallel circuits • compare parallel and series circuits in terms of voltage across components and current through them • identify uses of ammeters and voltmeters • explain why ammeters and voltmeters are connected differently in a circuit • explain why there are different circuits for lighting, heating and other appliances in a house	pages 123 123–9 109–10 110 133, 136	*Students:* • plan, choose equipment or resources for and perform first-hand investigations to gather data and use available evidence to compare measurements of current and voltage in series and parallel circuits in computer simulations or hands-on equipment • plan, choose equipment or resources for and perform a first-hand investigation to construct simple model household circuits using electrical components	pages 141–3 145

4. The amount of power is related to the rate at which energy is transformed

Students learn to: • explain that power is the rate at which energy is transformed from one form to another • identify the relationship between power, potential difference and current • identify that the total amount of energy used depends on the length of time the current is flowing and can be calculated using: Energy = VIt • explain why the kilowatt-hour is used to measure electrical energy consumption rather than the joule	pages 147 148 148 152–3	*Students:* • perform a first-hand investigation, gather information and use available evidence to demonstrate the relationship between current, voltage and power for a model 6 V to 12 V electric heating coil • solve problems and analyse information using $P = VI$ and Energy = VIt	pages 166 148, 163

5. Electric currents also produce magnetic fields and these fields are used in different devices in the home

Students learn to:	pages	*Students:*	pages
• describe the behaviour of the magnetic poles of bar magnets when they are brought close together	154–5	• plan, choose equipment or resources for and perform a first-hand investigation to build an electromagnet	169
• define the direction of the magnetic field at a point as the direction of force on a very small north magnetic pole when placed at that point	155–6	• perform a first-hand investigation to observe magnetic fields by mapping lines of force: – around a bar magnet – surrounding a straight DC current-carrying conductor – of a solenoid	167–9
• describe the magnetic field around pairs of magnetic poles	156–7		
• describe the production of a magnetic field by an electric current in a straight current-carrying conductor and describe how the right-hand grip rule can determine the direction of current and field lines	157–8	• present information using ⊗ and ⊙ to show the direction of a current and direction of a magnetic field	158, 164, 167–8
• compare the nature and generation of magnetic fields by solenoids and a bar magnet	158–60	• identify data sources, gather, process and analyse information to explain one application of magnetic fields in household appliances	161–2

6. Safety devices are important in household circuits

Students learn to:	pages		
• discuss the dangers of an electric shock from both a 240 volt AC mains supply and various DC voltages, from appliances, on the muscles of the body	131–3		
• describe the functions of circuit breakers, fuses earthing, double insulation and other safety devices in the home	133–5		

Module 3: MOVING ABOUT (chapters 9–11, pages 171–248)

1. Vehicles do not typically travel at a constant speed

Students learn to:	pages	*Students:*	pages
• identify that a typical journey involves speed changes	177	• plan, choose equipment or resources for and perform a first-hand investigation to measure the average speed of an object or a vehicle	194–5
• distinguish between the instantaneous and average speed of vehicles and other bodies	177, 182–4		
• distinguish between scalar and vector quantities in equations	180–1, 194	• solve problems and analyse information using the formula $v_{av} = \frac{\Delta \boldsymbol{r}}{\Delta t}$ where r = displacement	175–6, 184, 189–92
• compare instantaneous and average speed with instantaneous and average velocity	177, 182–6	• present information graphically of: – displacement vs time – velocity vs time for objects with uniform and non-uniform linear velocity	181–88, 194–5
• define average velocity as: $v_{av} = \frac{\Delta \boldsymbol{r}}{\Delta t}$	175		

2. An analysis of the external forces on vehicles helps to understand the effects of acceleration and deceleration

Students learn to:	pages	*Students:*	pages
• describe the motion of one body relative to another	175–7	• analyse the effects of external forces operating on a vehicle	218–22
• identify the usefulness of using vector diagrams to assist solving problems	173–4, 178–9	• gather first-hand information about different situations where acceleration is positive or negative	194–5
• explain the need for a net external force to act in order to change the velocity of an object	199–204	• plan, choose equipment or resources for and perform a first-hand investigation to demonstrate vector addition and subtraction	223–4
• describe the actions that must be taken for a vehicle to change direction, speed up and slow down	203	• solve problems using vector diagrams to determine resultant velocity, acceleration and force	219–20
• describe the typical effects of external forces on bodies including: – friction between surfaces – air resistance	202–5	• plan, choose equipment or resources for and perform first-hand investigations to gather data and use available evidence to show the relationship between force, mass and acceleration using suitable apparatus	224
• define average acceleration as $\boldsymbol{a}_{av} = \frac{\Delta \boldsymbol{v}}{\Delta t}$ therefore $\boldsymbol{a}_{av} = \frac{\boldsymbol{v} - \boldsymbol{u}}{t}$	178–9	• solve problems and analyse information using $\Sigma \boldsymbol{F} = m\boldsymbol{a}$ for a range of situations involving modes of transport	206–7, 215–17
• define the terms 'mass' and 'weight' with reference to the effects of gravity	197–8	• solve problems and amalyse information involving $F = \frac{mv^2}{r}$ for vehicles travelling around curves	220–2
• outline the forces involved in causing a change in the velocity of a vehicle when: – coasting with no pressure on the accelerator – pressing on the accelerator – pressing on the brakes – passing over an icy patch on the road – climbing and descending hills – following a curve in the road	202–5		
• interpret Newton's Second Law of Motion and relate it to the equation $\Sigma \boldsymbol{F} = m\boldsymbol{a}$	205–7		
• identify the net force in a wide variety of situations involving modes of transport and explain the consequences of the application of that net force in terms of Newton's Second Law of Motion	203–9, 213–14		

3. Moving vehicles have kinetic energy, and energy transformations are an important aspect in understanding motion

Students learn to:	pages	*Students:*	pages
• identify that a moving object possesses kinetic energy and that work done on that object can increase that energy	226–9	• solve problems and analyse information to determine the kinetic energy of a vehicle and the work done using the formulae: $E_k = \frac{1}{2}mv^2$ and $W = Fs$	227–9, 242–7
• describe the energy transformations that occur in collisions	230	• analyse information to trace the energy transfers and transformation in collisions leading to irreversible distortions	234–8, 242
• define the Law of Conservation of Energy	226, 229–30		

4. Change of momentum relates to the forces acting on the vehicle or the driver

Students learn to: • define momentum as $p = mv$ • define impulse as the product of force and time • explain why momentum is conserved in collisions in terms of Newton's Third Law of Motion	pages 231–2 231–3 240–1	*Students:* • solve problems and analyse secondary data using $p = mv$ and Impulse = Ft • perform first-hand investigations to gather data and analyse the change in momentum during collisions • solve problems that apply the principle of conservation of momentum to qualitatively and quantitatively describe the collision of a moving vehicle with: – a stationary vehicle – an immoveable object – another vehicle moving in the opposite direction – another vehicle moving in the same direction	pages 231–3, 240–1, 242–5 246–7 244–5

5. Safety devices are utilised to reduce the effects of changing momentum

Students learn to: • define the inertia of a vehicle as its tendency to remain in uniform motion or at rest • discuss reasons why Newton's First Law of Motion is not apparent in many real-world situations • assess the reasons for the introduction of low speed zones in built-up areas and the addition of airbags and crumple zones to vehicles with respect to the concepts of impulse and momentum • evaluate the effectiveness of some safety features of motor vehicles	pages 202 201–2 229, 235, 237–8 202–4, 234–8	*Students:* • gather and process first-hand data and/or secondary information to analyse the potential danger presented by loose objects in a vehicle • identify data sources, gather, process, analyse, present secondary information and use the available evidence to assess benefits of technologies for avoiding or reducing the effect of a collision	pages 236, 243, 246–7 229, 235, 237–8, 243, 245

Module 4: THE COSMIC ENGINE (chapters 12–14, pages 249–305)

1. Our Sun is just one star in the galaxy and ours is just one galaxy in the universe

Students learn to: • outline the historical development of models of the universe from the time of Aristotle to the time of Newton	pages 253–8	*Students:* • identify data sources, gather, process and analyse information to assess one of the models of the universe developed from the time of Aristotle to the time of Newton to identify limitations placed on the development of each model by the technology available at the time	pages 253–8, 268

2. The first minutes of the Universe released energy which changed to matter, forming stars and galaxies

Students learn to: • outline the discovery of the expansion of the universe by Hubble, following its earlier prediction by Friedmann • describe the transformation of radiation into matter which followed the 'Big Bang' • identify that Einstein described the equivalence of energy and mass • outline how the accretion of galaxies and stars occurred through: – expansion and cooling of the universe – subsequent loss of particle kinetic energy – gravitational attraction between particles – lumpiness of the gas cloud that then allows gravitational collapse	pages 260–62 264–5 258–60, 265 267	*Students:* • identify data sources and gather secondary information to describe the probable origins of the universe	pages 258–67, 268–9, 271–2

3. Stars have a limited life span and may explode to form supernovas

Students learn to: • define the relationship between the temperature of a body and the dominant wavelength of the radiation emitted from that body • identify that the surface temperature of a star is related to its colour • describe a Hertzsprung–Russell diagram as the graph of a star's luminosity against its colour or surface temperature • identify energy sources characteristic of each star group, including main sequence, red giants and white dwarfs	pages 276–7 276–7 277–8 279–82	*Students:* • gather secondary information to relate brightness of an object to its luminosity and distance • solve problems to apply the inverse square law of intensity of light to relate the brightness of a star to its luminosity and distance from the observer • process and analyse information using the Hertzsprung–Russell diagram to examine the variety of star groups, including main sequence, red giants and white dwarfs	pages 274–6, 285 274–6, 283 277–8, 280–2, 283–4

4. The Sun is a typical star, emitting electromagnetic radiation and particles that influence the Earth

Students learn to: • identify that energy may be released from the nuclei of atoms • describe the nature of emissions from the nuclei of atoms as radiation of alpha α and beta β particles and gamma γ rays in terms of: – ionising power – penetrating power – effect of magnetic field – effect of electric field • identify the nature of emissions reaching the Earth from the Sun • describe the particulate nature of the solar wind • outline the cyclic nature of sunspot activity and its impact on Earth through solar winds • describe sunspots as representing regions of strong magnetic activity and lower temperature	pages 289–90 288–9 292–3 292–3 295–7, 298–302 294	*Students:* • perform a first-hand investigation to gather information to compare the penetrating power of alpha, beta and gamma radiation in a range of materials • identify data sources, gather and process information and use available evidence to assess the effects of sunspot activity on the Earth's power grid and satellite communications	pages 305 299–301

ACKNOWLEDGEMENTS

The authors would like to thank the following people for their support during the writing of this book: Michael Andriessen gives special thanks to his wife, Christine, and sons, Sam and Luke, for their understanding and patience; Barry Mott is grateful to longstanding friend and colleague Bruce McKay for his patience and generous assistance; Ric Morante thanks the boys, who generously missed the park. Yoka McCallum's comments and advice are also appreciated, and thanks go to the staff of John Wiley for their attention to quality and detail.

The authors and publisher would like to thank the following copyright holders, organisations and individuals for their permission to reproduce copyright material in this book.

Images

- Austral International Press: **10.1** Austral International/ Pictor Uni Photo
- AAP Image: **5.13** AAP Image/Dean Lewins
- ANSTO: **14.3** reproduced with permission from the Australian Nuclear Science and Technology Organisation
- Carol Grabham: **8.29** © Paul Grabham
- Coo-ee Historical Picture Library: **11.2** (coach)
- Corbis Australia: **2.3** Corbis Australia/Zefa/Tony Latham; **11.1** Corbis Australia/Tim Wright; **12.6** Australian Picture Library/CORBIS/Geoffrey Clements; **12.7** (top right) Australian Picture Library/ CORBIS/Enzo R Paolo Ragazzini
- Corbis Royalty Free: **3.3(b)** and **11.2** (car) © Corbis Corporation; **3.4(c)** and **3.6** © Photodisc
- Coo-ee Picture Library: **5.10**
- David Malin Images: **12.3** © Anglo-Australian Observatory; **13.1** © Akira Fujii/David Malin Images
- Department of the Environment, Water, Heritage and the Arts: **8.5** reproduced by permission of the Department of the Environment, Water, Heritage and the Arts
- © Digital Vision: **3.8**, **6.1**, **9.1** and **14.20**
- Energy Australia: **8.4**
- Fairfax Photo Library: **8.1** Fairfax Photo Library/ Penny Stephens
- Getty Images: **3.4(a)** Getty Images/Stone/Raimund Koch; **10.18** Getty Images Sport/Clive Brunskill; **11.10** Getty Images/Stone
- HarperCollins Publishers UK: **13.7(b)** reprinted by permission of HarperCollins Publishers Ltd © Illingworth 1994
- Holden Ltd: **11.11**
- © imageaddict.com.au: **3.15(a)**, **4.9**, **4.19**, **6.47**, **6.50** and **8.26(b)**
- John Wiley & Sons Australia: **3.13** © John Wiley & Sons Australia/photo by Kari-Ann Tapp; **4.24** © John Wiley & Sons Australia/photo by Ron Ryan; **5.16** based on Department of Land & Water Conservation map; **7.1** John Wiley & Sons Australia, Ltd/photo by Claire Lord
- Loy Yang Power: **5.8**
- National Archives of Australia: **5.17** A 8746, KN13/8/ 75/62 — dam and hydro-electric plant, part of the Snowy Mountains Scheme
- NASA: **4.22**, **12.2**, **12.4**, **12.17**, **12.18**, **12.20**, **14.10** and **14.19** NASA; **12.1** NASA/STSCI; **14.5** NASA/ ISAS/ Yohkoh; **14.7** NASA/Goren Scharmer/Swedish Solar Vacuum Telescope; **14.8** and **14.9** NASA/Big Bear Solar Observatory/David Hathaway; **14.11** based on Marshall Space Flight Center and Science @ NASA's graph; **14.12** and **14.13** David Hathaway/NASA — Marshall Space Flight Center
- Newspix: **3.15(c)** Newspix; **5.1** Newspix/David Geraghty; **9.6** Newspix/Dean Marzolla; **11.6** Newspix/ Sam Ruttyn; **11.7** Newspix/Glenn Miller
- © Photodisc: **1.1**, **1.2**, **3.1**, **3.2(a)**, **3.3(a)**, **3.7**, **3.14**, **4.1**, **5.9**, **5.12**, **5.14**, **9.9**, **10.13**, **10.26(a)**, **11.2** (cyclist), **11.8**; **pages 1**, **75**, **171** and **249**
- photolibrary.com: **1.4** photolibrary.com/Science Photo Library/Martin Dohrn; **2.1**, **4.8** and **4.21** photolibrary.com/TEK image/Science Photo Library; **2.10** photolibrary.com/Matt Meadows; **3.4(b)**, **12.16** and **14.21** photolibrary.com/Science Photo Library/ NASA; **3.15(b)** photolibrary.com/Science Photo Library/Adam Hart-Davis; **4.42** photolibrary.com/ Science Photo Library/Cordelia Molloy; **5.2**, **5.3**, **5.4**, **5.5**, **5.7**, **8.16**, **12.9** (top right), **12.10** (both), **12.13**, **12.14** and **14.1** photolibrary.com/Science Photo Library; **5.6** photolibrary.com/Science Photo Library/ Sheila Terry; **7.24** photolibrary.com/Science Photo Library/Andrew Syred; **10.7** photolibrary.com/Science Photo Library/Bill Sanderson; **11.2** (train) photolibrary.com/Chris Jones; **12.7** (top left), **12.8** (bottom right) and **12.11 (a)** photolibrary.com/ Science Photo Library/Dr Jeremy Burgess; **12.8** (bottom left) photolibrary.com/Science Photo Library/Library of Congress; **12.9** (top left) photolibrary.com/Science Photo Library/Jen-Loup Charmet; **12.12** photolibrary.com/Science Photo Library/George Bernard
- Réunion des Museés Nationaux: **6.3** © Photo RMN — Gérard Blot
- Dr Ragbir Bhathal: **13.7(a)** Bhathal, R, *Astronomy for the HSC*, Kangaroo Press 1993. Reproduced by permission of Ragbir Bhathal
- © Russell Kightley Media: **12.22**
- Scala: **12.11(b)** Galileo Galilei (1564–1642): Galileo 48 c 28 r. Florence, Biblioteca Nazionale. Photo Scala, Florence, courtesy of the Ministero Beni e Att. Culturali

- Snowy–Hydro Limited: **5.15** copyright Snowy Hydro Limited. Only to be reproduced with the prior permission of Snowy Hydro Limited
- Space Environment Center: **14.17** and **14.18** Space Environment Center, Boulder CO, National Oceanic and Atmospheric Administration, US Dept of Commerce

Text

- *Physics Stage 6 Syllabus* © Board of Studies NSW, for and on behalf of the Crown in right of the State of New South Wales, 2002 (syllabus grid, pages **xi–xiv**)
- Data and Formulae sheets from *Physics Higher School Certificate Examination* © Board of Studies NSW, 2002 (formulae and data sheet, page **309**)
- Key words from: The New Higher School Certificate Support Document © Board of Studies, 1999. Most up-to-date version available at www.boardofstudies.nsw.edu.au/syllabus_hsc/glossary_keywords.html (key words list, page **311**)

Every effort has been made to trace the ownership of copyright material. Information that will enable the publisher to rectify any error or omission in subsequent editions will be welcome. In such cases, please contact the Permissions Section of John Wiley & Sons Australia, Ltd.

THE WORLD COMMUNICATES

CHAPTER

1 WAVES: MOVERS OF ENERGY

Figure 1.1 *Mobile phones are now established as one of the fastest growing methods of communication.*

Remember

Before beginning this chapter, you should be able to:

- identify that waves carry energy
- recall that the properties of waves include frequency, wavelength and speed.

Key content

At the end of this chapter you should be able to:

- describe waves as a way that energy is transferred which, depending on the wave type and medium, can occur in one, two or three dimensions
- define the wave model and apply the terms specific to its definition: medium, displacement, amplitude, period, compression, rarefaction, crest, trough, transverse waves, longitudinal waves, frequency, wavelength, velocity
- represent diagrammatically the troughs and crests of a transverse wave and calculate the wavelength and amplitude
- relate particle motion to the direction of energy propagation in transverse and longitudinal waves
- present diagrams of transverse and longitudinal waves, showing wavefronts and direction of particle motion and propagation
- explain that mechanical waves need a medium in which to propagate whereas electromagnetic waves do not
- gather information about the transmission of waves in slinky springs, water waves and ropes
- use displacement–time graphs to analyse transverse wave motion
- recognise the relationship between the velocity, frequency and wavelength of a wave and solve problems using $v = f\lambda$
- describe the energy transformations that take place when using a mobile telephone.

Our society is in the 'information age' and relies on messages carried as energy pulses by waves. Waves are carriers of energy and are used increasingly as carriers of messages for communication. We use a broad range of wave types to transfer messages by many different means, whether it is by speaking to the person next to us, or by talking into a phone or typing an email. Mass communication has become possible through the media of television and the internet.

In this chapter we examine some of the basic properties of waves. In chapters 2 and 3 we will look at the properties of sound waves and electromagnetic waves and how we harness those waves to communicate.

1.1 THE WAVE MODEL

Most people understand the term 'wave' by visualising the surface appearance and motion of water — for example, the ripples on a pond or the swells we see in the deep ocean. We are comfortable with the concept of a wave we can see.

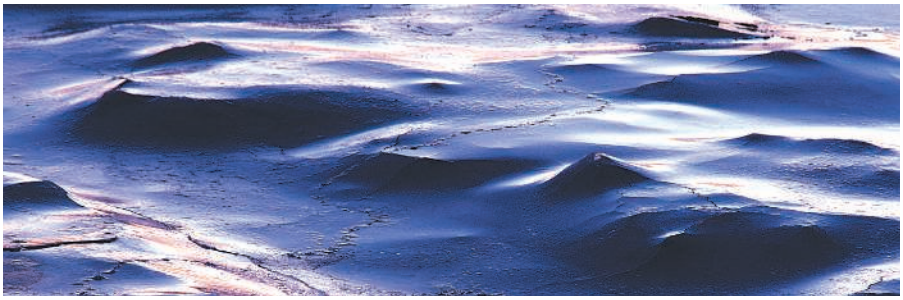


Figure 1.2 ***The image of waves on the ocean helps us to visualise the properties of other types of waves that we cannot see.***

Other phenomena that we cannot see can also be explained as wave motions. The wave model enables us to visualise and describe everyday phenomena that have wave properties. With the wave model, we take information we can see or identify — such as light, sound and electricity — and visualise it as wavelike. This helps us to describe and explain its behaviour.

A sine wave is the curve that results when a plot is made of $y = \sin x$.

In cross-section, waves such as ripples or swells have the shape of a sine curve or **sine wave**, as shown in figure 1.3.

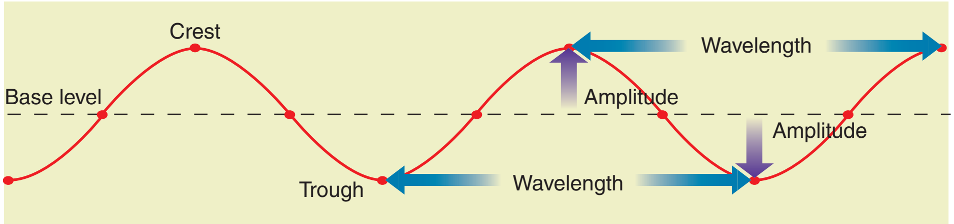


Figure 1.3 ***Wavelength, amplitude, crest and trough are useful terms in describing a wave.***

Using the concept of a wave as a sine wave, we can name and measure a number of wave features.

- *Crests* are the highest points of the waves.
- *Troughs* are the lowest points of the waves.
- The *wavefront* is the face of the wave as it propagates forward. It can be a trough or crest and it is perpendicular to the direction of wave travel.
- Two points on a wave are in *phase* if, at a particular instant, they have the same *displacement* and the same velocity.
- The *amplitude* is the maximum size of the particle displacement from the undisturbed state.

- The *wavelength* is the distance between two adjacent crests or troughs in a wave and is assigned a symbol, λ, when used in equations. A wavelength is the length of a wave that includes an entire trough and an entire crest. A full wavelength of a sine wave is drawn when the graph of $y = \sin x$ is plotted for values of x from 0 to 360°.
- The *frequency* is the number of waves that pass a fixed point per second. The frequency is assigned a symbol, f, when used in equations. The frequency of waves is usually measured in cycles per second, or hertz (Hz). One hertz is one cycle or wavelength passing a point per second.
- The *period* is the time it takes a single wave to pass a fixed point. It is assigned a symbol, T, when used in equations. The period and the frequency are related through a reciprocal relationship. That is: $f = \frac{1}{T}$ or $T = \frac{1}{f}$.

To propagate is to transmit through space or a medium.

- The speed or *velocity* at which a wave **propagates** is how fast the wave transfers energy away from a source. If the wavelength and frequency of a wave are known, we can calculate the velocity. For example, if a wave has a wavelength of 10 cm and a frequency of 5.0 Hz, then the wave must travel $10 \times 5.0 = 50$ cm in a second; that is, its velocity is 50 cm per second. Velocity, v, is the product of the wave's frequency, f, and wavelength, λ; that is $v = f\lambda$.
- The *medium* is the material in which the wave is propagating. Electromagnetic waves do not require a medium.

SAMPLE PROBLEM 1.1

Calculating wave velocity

An off-shore swell consists of waves with a wavelength of 10 m. The frequency of those waves passing a fixed point was measured at 0.5 Hz. What is the velocity of the wave motion?

SOLUTION

$$\begin{aligned} v &= f\lambda \\ &= 0.5 \times 10 \\ &= 5 \text{ m s}^{-1} \end{aligned}$$

The wave is travelling at 5 m s^{-1}.

SAMPLE PROBLEM 1.2

Calculating wave frequency

Light, an electromagnetic wave, travels at 3.00×10^8 m s^{-1}. What is the frequency of green light of wavelength 550 nm (550×10^{-9} m)?

SOLUTION

$$v = f\lambda$$

$$\begin{aligned} f &= \frac{v}{\lambda} \\ &= \frac{3.00 \times 10^8}{550 \times 10^{-9}} \\ &= 5.45 \times 10^{14} \text{ Hz} \end{aligned}$$

SAMPLE PROBLEM 1.3

Sound waves travel at 340 ms^{-1} in air. The frequency of sound coming from an ultrasound (above hearing range) motion detector is 40 000 Hz. Calculate the wavelength of the sound waves produced by the detector.

SOLUTION

$$v = f\lambda$$

$$\begin{aligned} \lambda &= \frac{v}{f} \\ &= \frac{340}{40\,000} \\ &= 8.5 \times 10^{-3} \text{ m} \end{aligned}$$

1.2 WAVES TRANSFER ENERGY

The source of all waves is a vibration. The energy of that vibration passing from the place of origin to a place further away is the wave. For example:

- Figure 1.4 shows water waves produced by a pebble thrown into a still pond. The pebble sets the water particles vibrating in an up and down motion that moves out from the source of disturbance. Some of the movement energy of the pebble has been transferred to the water particles.
- Vibrating vocal cords in the larynx produce sound waves in the form of the human voice (sound waves are discussed in chapter 2).
- A vibrating electric current in an aerial produces radio waves (see chapter 3, page 30).

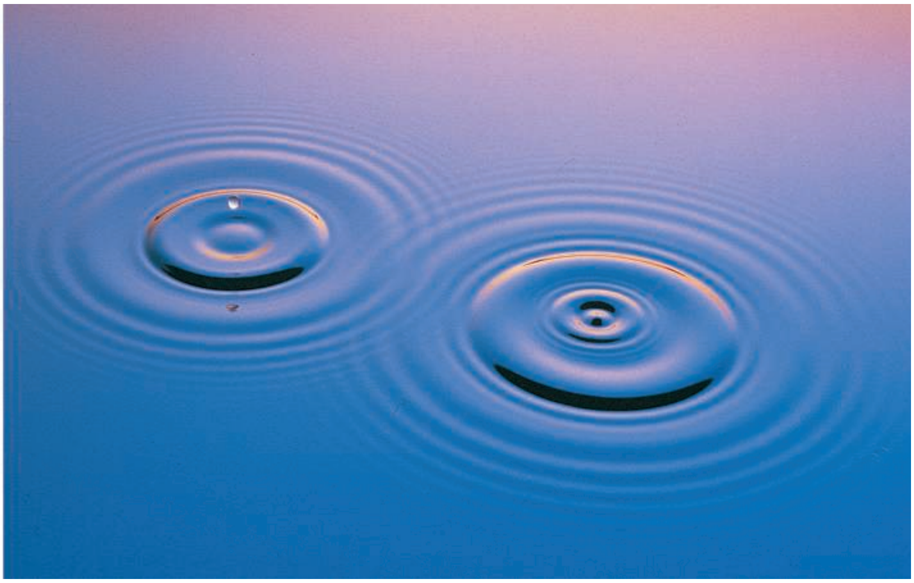


***Figure 1.4** Ripples on a still pond. Energy is carried by the wave away from a source of disturbance, such as a pebble thrown into the pond. The material of the medium through which the wave passes is not carried along — it is energy, not matter, that is transferred from one place to another.*

For all wave motions, whatever their origin, the transfer of energy is in the direction in which the wave is travelling. That transfer of energy is always away from the source of vibration. When the source of the waves is from a point acting as a source of vibration (called a point source) the waves radiate out from that point.

Depending upon the type of wave and the medium in which they are travelling, waves may travel in one, two or three dimensions.

PRACTICAL ACTIVITIES

1.1

Investigating waves in a slinky spring

An example of a wave travelling in *one dimension* is the motion of either a transverse or longitudinal wave in a slinky, or a transverse wave travelling along a rope. In this case the medium confines the wave to the rope or slinky (see figure 1.5). The energy of the wave motion has only one dimension in which to travel.

***Figure 1.5** In a wave that is made to travel along a rope, the wave is confined to the rope as the medium so it only travels in one dimension.*

An example of a wave travelling in *two dimensions* is a transverse wave travelling from a point source of disturbance in still water. A pebble thrown into a still, flat-bottomed pond will produce a wave travelling outwards with a circular wavefront away from the initial disturbance (see figure 1.6).

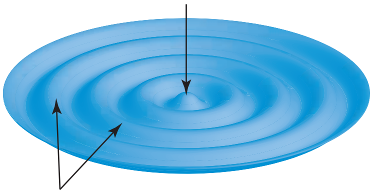


***Figure 1.6** A circular wavefront propagating from a point source. Note that the further the wave is from the source the straighter the wavefront appears because the wavefront represents a smaller arc of the circle as it expands. This implies that the energy in a constant area, or length of wavefront, is lower the further the wavefront is from the source.*

PRACTICAL ACTIVITIES

1.2

Observing water waves

As an example of waves travelling in *three dimensions*, consider a point source of sound — it results in a sound wave that immediately travels away from the source in three dimensions with a spherical wavefront. Similarly, a point source of light will illuminate a three-dimensional space.

We know that waves carry energy because waves have the ability to transfer their energy to other bodies. In doing so, the waves do work on those bodies. Waves transfer energy without the transfer of matter.

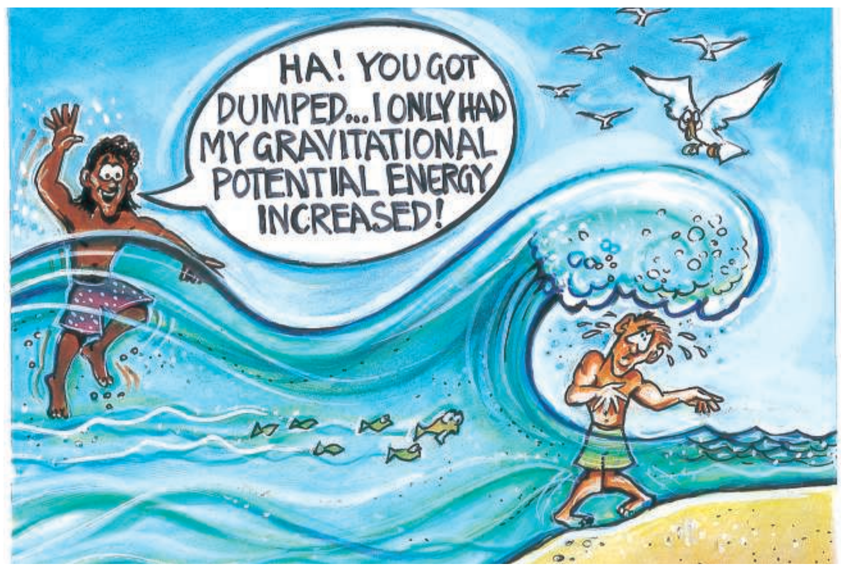


Figure 1.7 *Surfers receive energy from water waves. They gain potential energy as they are lifted toward the crest of a wave. This is turned into kinetic energy as they fall down the face of the wave and are pushed forward by the motion of the wave.*

This transfer of energy by waves can be demonstrated when:

- a surfer bobs up and down in the swell (see figure 1.7 above)
- the free end of a rope tied to a post is given a jerk
- a slinky spring lying flat on a table, slightly stretched and held firmly at one end, is given a sideways shake.

In each case a disturbance travels through the medium (the water, the material of the rope, the slinky) but the medium does not move forward with the disturbance. It is the disturbance that travels through each of the materials that forms the waves. Waves pass from one place to another without taking material with them. They are simply a means of transferring energy from place to place. Wave motion is the most important means of transferring energy from place to place within the universe.

PHYSICS IN FOCUS

Energy transfer — see it, hear it and feel it!

Energy transfer by waves is occurring all around us.

- Light waves travel from the Sun and provide the energy for photosynthesis and hence, life on Earth.
- Vision is the result of energy transfer. We see because light waves that reach our eyes stimulate a response in nerve endings in the retina. Different colours or intensities of light have different amounts of energy and so stimulate different responses in the nerves. This stimulation information is relayed to the brain as nerve impulses representing the picture and is what we see.
- Infra-red waves from hot objects such as the Sun or a radiator are detected by nerve endings on our bodies. We absorb them and become warm ourselves.
- Microwaves are absorbed by water molecules causing the water molecules to gain kinetic energy. This energy cooks our food.
- Sound waves carry the energy to vibrate our eardrums, allowing us, for example, to communicate or listen to our favourite music.
- Earthquake waves transfer stored energy from the rocks of the Earth's crust as a shake, often with catastrophic results.

1.3 PROPAGATION OF WAVES

A **mechanical wave** is a wave that requires the movement of particles to propagate forward.

An **electromagnetic wave** is a wave that propagates as a perpendicular electric and magnetic field. Electromagnetic waves do not require a medium for propagation.

An **oscillation** is a vibration about a fixed or equilibrium point.

Waves are categorised according to how they propagate or transfer energy from place to place. There are two major groups of waves: **mechanical waves** and **electromagnetic waves**.

Mechanical waves involve the transfer of energy through a medium by the motion of particles of the medium itself. The particles move as **oscillations** or vibrations around a fixed point. After the wave has passed, the particles that move are in exactly the same place as before they were disturbed. There is no bulk transfer of particles from one place to another. This is exactly the case with water waves away from the shore of the ocean or pond.

Electromagnetic waves, such as radio and light, are able to propagate through the near vacuum of space without a medium. These waves do not need the movement of any particles to propagate (as mechanical waves do), and they are not subject to the same energy losses due to friction between particles. Therefore, they potentially have much greater travel ranges. This is vital for long-distance communication. Electromagnetic waves are discussed in chapter 3.

Because mechanical waves transfer energy as an oscillation of particles, it follows that there must be a material substance to act as a medium for the transmission of a mechanical wave.

Mechanical waves are classified as either transverse or longitudinal according to the direction of disturbance or vibration relative to the direction of energy flow through a material.

- In a transverse wave, the particles of the medium vibrate in a plane that is perpendicular to the direction of propagation of the wave.
- In a longitudinal wave, the particles of the medium vibrate in the same direction as the direction of propagation of the wave.

Modelling a transverse wave

The vibrational component of a transverse wave involves particles undergoing a motion perpendicular to the direction of propagation. An example of a visible transverse wave are the ripples produced by a disturbance in a pond. The motion of the water particles is perpendicular to the direction of wave travel at any point on the wavefront.

Transverse waves can be modelled with a slinky spring. To do this, one end of the slinky should be tied to a post and the slinky should be stretched. If the free end of the stretched slinky is moved up and down or left and right with sharp jerks, a transverse wave should form in the spring. The energy of the pulse produced by the shaking should travel forward as a sine wave form. If this activity is done correctly, a pulse of vertically or horizontally displaced slinky coils travels along the length of the slinky as shown in figure 1.8. If you jerk the spring regularly, it should be possible to create a wave form with the slinky that appears to remain still even though you are continuing to jerk the spring (and hence putting energy continuously into the spring movement). This wave that appears to be standing still is called a standing wave. The motion of the spring coils is at ninety degrees to the direction of wave travel and is about their equilibrium position. The energy pulse of the wave propagates forward. To increase the amplitude of the standing wave you must shake with more energy but with the same frequency.

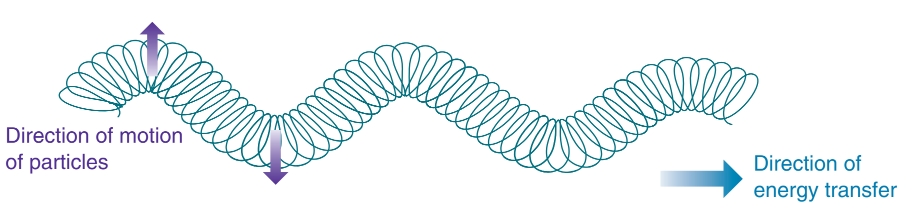


Figure 1.8 *A transverse wave in a slinky spring*

Electromagnetic waves are transverse waves because they consist of alternating electric and magnetic force fields at ninety degrees to one another and the direction of energy propagation and transfer. This is shown in figure 1.9 and is discussed further in chapter 3.

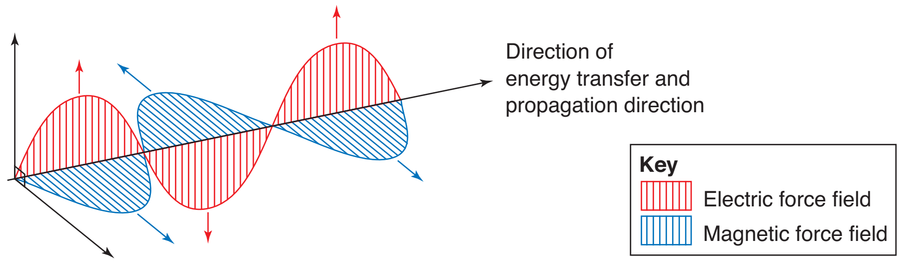


Figure 1.9 ***A representation of an electromagnetic wave force field showing the energy propagation direction***

Modelling a longitudinal wave

Sound is one of the most common longitudinal waves. You can't see sound, although you can feel it as a vibration by touching a speaker or talking into an inflated balloon and feeling the other side of the balloon.

Sound waves are discussed in detail in chapter 2.

Modelling a longitudinal wave can also be done with a slinky spring. To do this, tie one end of the slinky to a post with a string. Stretch the slinky out and bunch up five or six coils from the ends of the slinky and press them together with your fingers. You are storing energy in the compressed spring coils. If you let the compressed coils go, you will see a pulse of compressed coils travel along the length of the slinky as shown in figure 1.10. The pulse produced in the spring acts just like a longitudinal wave. The *compression*, or wave pulse, travels along the length of the spring. On each side of the compressed pulse is a zone where the slinky coils are spread apart. These areas are known as *rarefactions*. The motion of the particles produced by the moving energy of the wave is back and forth in the direction of wave propagation. In the case of sound waves, the pulses are compressed air particles (compressions) surrounded on either side by air particles that have been spread apart (rarefactions). Compressions are zones of high air pressure and rarefactions are zones of low air pressure.

See page 16 for an explanation of compressions and rarefactions in relation to sound waves.

The wavelength, λ, of a longitudinal wave is equivalent to the distance between adjacent zones of compression or rarefaction (see figure 1.10). The amplitude of a longitudinal wave is equal to the size of the maximum displacement of the particles from their equilibrium position.

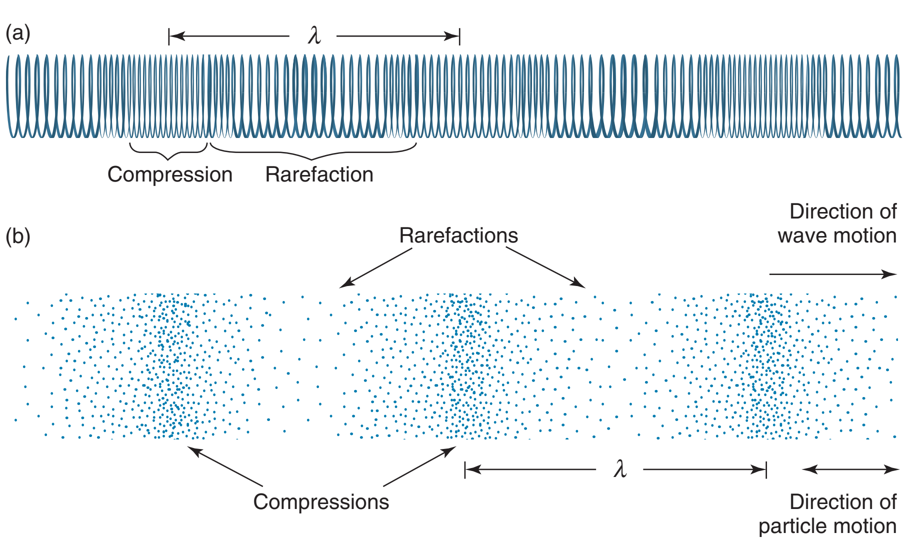


Figure 1.10 ***Longitudinal wave in (a) a slinky and (b) air***

PHYSICS IN FOCUS

Energy transformations in a mobile phone call

Mobile phones have a built-in microphone that changes sound waves into electrical signals. These electrical signals are digitised (converted into a code of 1s and 0s) and transmitted as radio waves to a base station. The base station consists of a system of antennae on top of a tower or tall building. Each base station accepts and can transmit radio signals from three adjacent hexagonal-shaped areas called cells. The arrangement of base station aerials in cells is shown in figure 1.11.

Each base station is connected to a switching centre by a cable network that carries the signal as electrical impulses. The impulses have been produced by radio-wave energy interacting with the aerial. Each switching centre is connected to other switching centres and base stations. There are three main possible paths for the signal to take.

1. If the telephone call occurs between a mobile and a distant fixed telephone, the signal may be converted into light and travel along an optical fibre network to a distant switching centre close to its destination. From the switching centre, it is moved into the copper-wire network as an electrical impulse and is decoded in a receiving telephone.
2. If the telephone call occurs between a mobile and a fixed telephone close to the aerial, it is converted into an electrical impulse in a copper wire. It may remain as an electrical impulse in the copper-wire network until it reaches a switching centre close to its destination.

   As in the previous case, the signal is then routed as an electrical impulse along copper wire to a receiving telephone where the signal is converted back into sound energy.
3. If the telephone call is from one mobile to another mobile, the signal will be transferred to a switching centre close to a base station servicing the cell near where the receiving mobile telephone is located. The signal from the switching station is fed to the base station as an electrical signal and broadcast as radio waves to the mobile phone. Once at the mobile the radio signal is converted back into electrical impulses. The electrical impulse signal is then converted by the speaker in the phone into sound.

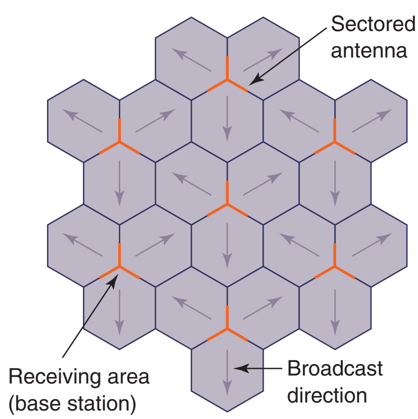


Figure 1.11 *The arrangement of base stations servicing mobile phone cells*

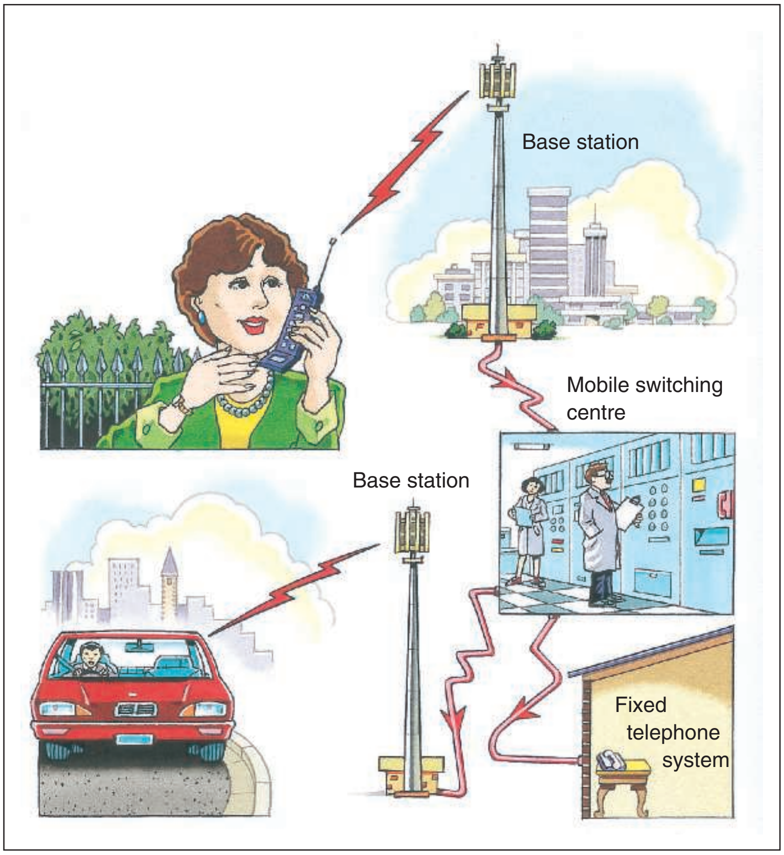


Figure 1.12 *Some of the pathways followed by a mobile phone call*

1.4 GRAPHING WAVE MOTION

1.3

Relating frequency and amplitude

The displacement of a point, as a wave passes, can be plotted against time. This produces a graphical representation of the wave motion (providing the wave is not a standing wave). With a transverse wave this means simply plotting the relative displacement of the point perpendicular to the wave propagation direction. If the wave is of constant frequency and amplitude, the shape of the graph (shown in figure 1.13) will be that of the curve $y = n \sin ft$, where n is the amplitude of the wave and f is the variable that determines the frequency. If this equation represents a displacement–time graph of the wave motion, the variable y represents the displacement, whereas sin t represents time. This relationship can be explored simply using a graphics calculator or many of the function drawing or spreadsheet programs available for computers.

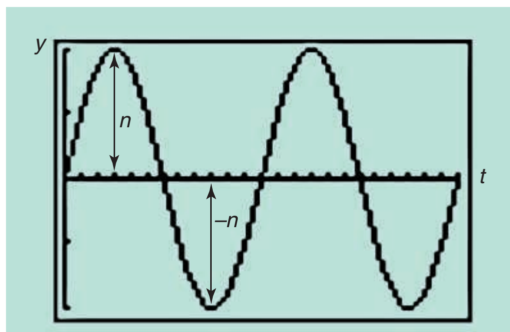


Figure 1.13 ***Sine waves can be generated using computer graphing programs.***

CHAPTER REVIEW

SUMMARY

- Waves can be categorised as:
  – mechanical waves, consisting of particles with energy, which require a medium for propagation
  – electromagnetic waves, which do not require a medium for propagation.
- A wave consists of two motions:
  (1) a uniform motion in the direction of wave travel; this is the direction of energy transfer.
  (2) a vibration of particles or fields about an equilibrium or central point.
- The vibration disturbance component of the wave may occur:
  – at right angles (90°) to the direction of propagation; these waves are called transverse waves.
  – in the same direction as the direction of propagation; these waves are called longitudinal waves.
- For transverse and longitudinal waves, $v = f\lambda$ and $f = \frac{1}{T}$.
- Waves transmit energy but do not transfer matter.

QUESTIONS

1. Describe how each of the following observations allows you to determine that waves are carriers of energy.
   (a) On a camping trip, infra-red and visible radiation from the Sun is absorbed by a solar shower, heating the water.
   (b) Sound waves hitting the diaphragm of a microphone cause the still diaphragm to begin to vibrate.
   (c) A tsunami destroys a coastal village.
   (d) A big surf removes the sand from a 10 km stretch of beach.
   (e) Light falling on a photovoltaic cell produces stored chemical energy in a battery that is used to produce electrical energy to power a solar garden light.
2. Describe each of the following waves as propagating in one, two or three dimensions:
   (a) the light emitted from the Sun
   (b) sound from a bell
   (c) a sound wave travelling along a string telephone made from a tight string and two tin cans
   (d) a water wave produced by dropping a rock into the centre of a lake
   (e) a compression wave produced in a slinky.
3. Define the following terms as they would apply when describing the wave model:
   (a) medium
   (b) displacement
   (c) amplitude
   (d) period
   (e) compression
   (f) rarefaction
   (g) crest
   (h) trough.
4. Draw a representation of a transverse wave, showing displacement versus time as a sine wave and label the following features:
   (a) the amplitude of the wave
   (b) a change in frequency of the wave
   (c) a wavefront
   (d) a trough
   (e) a crest.
5. In movies it is common to see a spacecraft blown up, accompanied by a large bang. With reference to the properties of mechanical waves, explain why this is impossible in space.
6. An astronomer tells you that observing the star she is showing you is like looking back in time 100 million years. Identify what this tells you about the light from the star and its nature of travel.
7. Light travels at a velocity of 3×10^8 m s^{-1}. If the light reaching Earth from a blue star has a wavelength of 410 nm (410×10^{-9} m), what is the frequency of the light?
8. Look at the transverse wave represented in figure 1.14. Calculate the frequency of the wave.

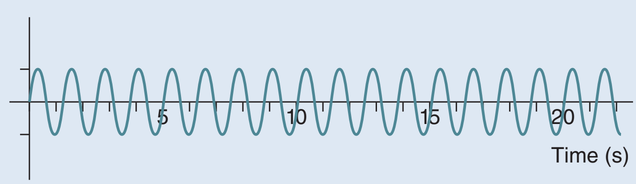


Figure 1.14

9. A cathode-ray oscilloscope (CRO) is a device that enables you to look at the electrical signal produced by a sound wave hitting the diaphragm of a microphone. The CRO acts like a sensitive voltmeter. Identify which property of the sound wave produces the sympathetic fluctuations in the voltage generated in the microphone.

10. The CRO trace in figure 1.15 is produced by a sound wave. The time base of the CRO is set at a constant value. That means every horizontal division of the figure represents a constant 0.001 s. Is the velocity of this wave constant? Explain your answer.

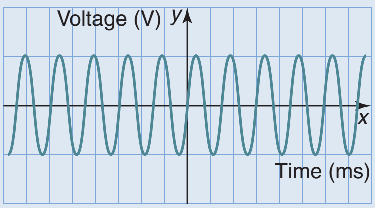


Figure 1.15 ***The CRO trace of a sound wave***

11. Sound travels in air at a speed of 330 m s^{-1}. Calculate the wavelength of a sound wave with a frequency of 256 Hz.
12. The centre of a compression in a sound wave (longitudinal or compression wave) is equivalent to the crest of a transverse wave, and the centre of a rarefaction is equivalent to the trough of a transverse wave. Use this information to present, in a diagram, a transverse wave representation of the sound wave shown in figure 1.16. Lines close together represent high pressure zones (compressions) and lines spread far apart (rarefactions) represent low pressure zones.

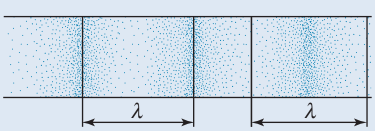


Figure 1.16 ***A representation of a sound wave***

13. Explain why both representations of a wavelength λ on figure 1.16 are correct.
14. Sound is travelling in a medium at 330 m s^{-1}. Using lines to represent wavefronts, present a scaled, labelled diagram to represent a 100 Hz sound wave. Include two compressions and two rarefactions in your drawing.
15. A p-type earthquake wave is a longitudinal wave, whereas an s-type earthquake wave is a transverse wave. Describe how each of these wave types would express itself in terms of earth movements under foot as it passed. Assume that the waves you are comparing are of equal intensity and that the waves are travelling along the ground surface towards you.
16. Identify the features of sound that are wave-like.
17. Explain why it is necessary to use a wave model to explain features of the behaviour of sound and light.
18. Light is an electromagnetic wave that is considered to be transverse in nature. What feature of this wave type suggests it is a transverse wave?
19. Explain why it is not possible to have a mechanical wave in a vacuum.
20. On a CRO trace of a sound wave that looks like figure 1.15, explain what the base line represents in terms of the sound wave.
21. During a mobile telephone conversation, a number of different energy transformations take place. Identify the transformations that involve mechanical waves and those that involve electromagnetic waves. Use figure 1.12 (page 9) to assist you in identifying these transformations.

CHAPTER REVIEW

1.1 INVESTIGATING WAVES IN A SLINKY SPRING

Aim

To observe and investigate the behaviour of waves (or pulses) travelling along a slinky spring

Apparatus

slinky spring
other pieces of apparatus as required

Theory

In a slinky spring, the velocity of the wave is quite small and it is not difficult to make observations of waves or pulses as they move along the spring. Many important properties of waves can be observed using this simple equipment.

Method

1. Stretch out a slinky spring, preferably on a smooth floor, until it is about three metres long. (Clamp one end to a fixed object.)
2. Displace the end of the spring to one side and quickly return it to its original or equilibrium position. This should cause a transverse pulse to travel along the spring.
3. Now, instead of a movement to the side to send a transverse pulse, gather up a few coils of the spring and then release them. This should produce a longitudinal or compression pulse.
4. You are now in the position to make investigations of the behaviour of pulses travelling in springs. It will be better to use transverse rather than longitudinal pulses for your experiments. You should be able to devise a series of simple experiments that can be used to investigate the following.
   (a) Do the pulses you produce really carry energy?
   Design and demonstrate a simple experiment that will show that the pulses do carry energy.
   (b) Do the pulses lose energy as they travel along?
   What observations can you make that show that the energy carried by each pulse is slowing dying away? Suggest a reason why this is happening. You may be able to think of a way to reduce the rate at which the energy is lost. This will probably involve changing the condition of the slinky.
   (c) Does the speed of a pulse depend on the condition of the slinky?
   As the tension of the slinky is changed, what happens to the speed of a pulse?
   (d) Does the speed of a pulse depend on the amplitude of the pulse?
   Is it possible to detect a change in the speed of the pulse along the slinky as the amplitude of the pulse is changed?
   (e) Does the speed of a pulse depend on the length of a pulse?
   Is it possible to detect a change in speed of the pulse along the slinky as the length of the pulse is changed?
   (f) Do identical pulses travel at the same speed along identical slinkies?
   Obtain a second slinky and lay the two side by side. Investigate how identical pulses travel in what are hopefully identical slinkies. The two ‘free’ ends of the slinkies could be fixed to a metre ruler (or wooden board) and movement of this could launch identical pulses simultaneously in the two slinkies.
   (g) Do identical pulses travel at the same speed along slinkies that are at a different tension?
   Changing the tension of one slinky by clamping a few coils to the board could enable you to observe identical pulses travelling in different media.

Extension work

(h) *Reflection from a fixed end*
You could also investigate what happens to pulses when they are reflected from the clamped end of a slinky.

(i) *Reflection from a free end*
You might repeat investigation (h) with the end tied to a length of light string so that it is free to move when the pulse reaches the end of the slinky. There will still be a reflected pulse but it should be different from the reflected pulse in (h).

(j) *Crossing boundaries*
What happens when a pulse travels from one spring to another?
If you have a different spring, connect it to a slinky and see what happens when a pulse in one spring travels into the other.

1.2 OBSERVING WATER WAVES

Aim

To observe a wave motion travelling in two dimensions

Apparatus

20 corks
small tank of water or a shallow, still-water pond

Theory

A simple transverse water wave is a wave travelling in two dimensions.

Method

1. Place 20 corks in a ring in the water.
2. Drop a small mass such as a stone in the centre of the ring of corks. This will make a wave in the water emanating from where the stone landed.
3. Observe what happens to all of the corks.

Analysis

Any movement of the corks outwards from the central disturbance is minor and at a very slow rate compared with the rate of energy transfer as indicated by the wavefront travelling away from the source.

Questions

1. In how many dimensions does the wave propagate?
2. What does this show about the energy of the wave motion from the central disturbance point?
3. In the previous practical activity, you observed that the energy carried by a pulse in the slinky was gradually lost. The same thing happens with the water waves. Compare the reasons for the decrease in the amplitude, and hence energy of a wave, in the slinky and in a wave spreading out on water.

1.3 RELATING FREQUENCY AND AMPLITUDE

Aim

To explore the relationship between the displacement and time of constant frequency waves with varying amplitude described by the equation:

$y = n \sin ft$

where

y = displacement of the wave
n = amplitude of the wave
f = frequency of the wave
t = time

Apparatus

access to a graphics calculator or a graphing program for the computer. Some graphing programs can be downloaded from the internet.
access to a printer

Method

1. Plot the equation given under 'Aim' above into a graphics calculator or a graphing program on the computer.
2. Plot graphs with the following variables and if a printer is available, print the graphs out.

FREQUENCY (Hz)	AMPLITUDE (units)
1	1
2	2
0.5	1
4	1
1	4

Analysis

Study the graphs to ensure that you can identify the features of amplitude and frequency.

CHAPTER 2

SOUND IS A WAVE

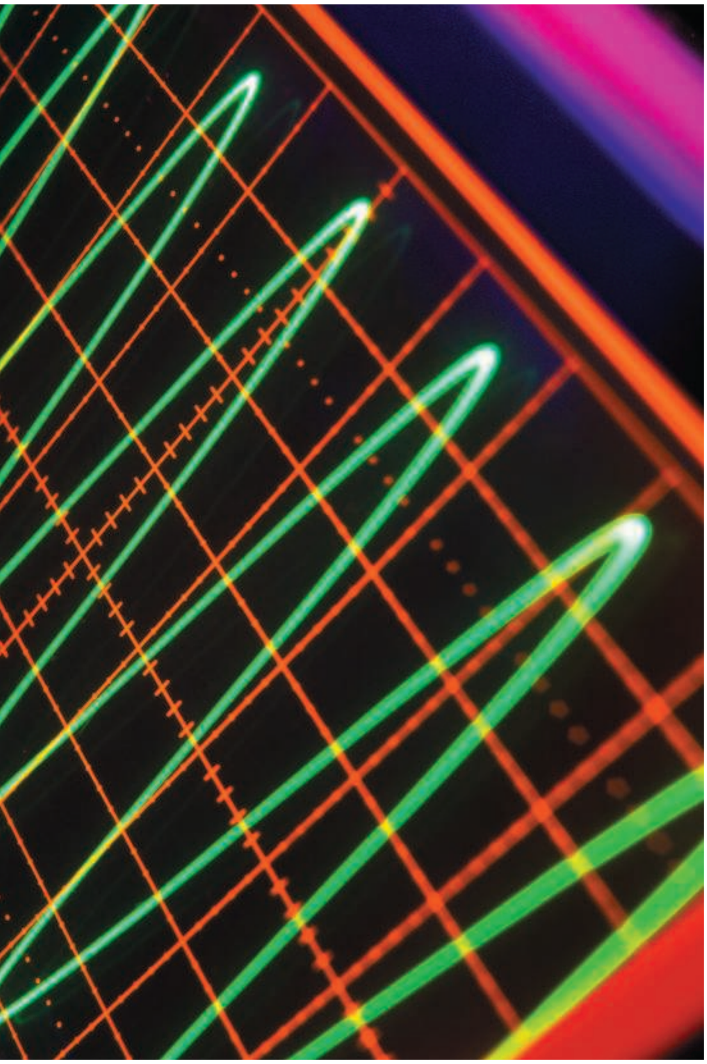


Figure 2.1 *A cathode-ray oscilloscope (CRO) is an electronic device that can be used to study sound waves. The CRO enables you to see sound waves.*

Remember

Before beginning this chapter, you should be able to:

- describe sound as a mechanical, longitudinal wave and describe the characteristics of sound.

Key content

At the end of this chapter you should be able to:

- understand that sound waves are the vibrations or oscillations of the particles of a medium
- relate compressions and rarefactions of sound waves to the crests and troughs of a transverse wave representation from a cathode-ray oscilloscope (CRO)
- perform an investigation that enables you to gather information to determine the relationship between the frequency and wavelength of a sound wave in a medium of constant properties
- use a CRO to gather information about the frequency, amplitude and velocity of sound waves and to observe and analyse the different sources of sound waves
- explain that pitch is related to the frequency of a sound wave, and volume is related to the amplitude of a sound wave
- explain that an echo is the reflection of a sound wave
- identify the conditions necessary to hear an echo
- describe superposition and present graphical information showing the superposition of waves
- use a CRO to observe the superposition of two sound waves travelling in air.

Throughout this chapter much emphasis will be placed on the use of the cathode-ray oscilloscope (CRO) or alternative computer technology to see sound waves. Its use is a mandatory part of your study of sound.

Sound is everywhere in our lives. Hearing is one of our dominant senses and people who lose the ability to hear usually take significant steps to restore it; for example, with hearing aids such as the bionic ear or cochlear implant. Sound is all around us, yet we often fail to notice it. Stop for a few seconds and listen to the sounds you can hear. Your sound-wave detector, the ear, is extremely sensitive. Most young people have the ability to hear sounds within the frequency range 20 Hz to around 20 000 Hz. Unfortunately, as you get older this range compresses as the ear loses sensitivity.

2.1 SOUND: VIBRATIONS IN A MEDIUM

All sound waves are vibrations in a medium that result in pressure variations within that medium. In air, the presence of those vibrations you know as sound is relatively easy to detect. You can see the vibrations and their effect by placing a piece of paper in front of your mouth in contact with your lips and speaking loudly. You will feel the paper vibrate because of the pressure differences in the vibrating air or sound wave. Alternatively, place a piece of paper in front of a stereo speaker, turn up the volume and observe and feel the paper vibrating. Sound pressure waves can blow a person off their feet if the sound is loud enough (sufficiently high in amplitude).

The origin of a sound wave in any medium is always a vibration. The frequency of the original vibration determines the frequency of the sound produced by that vibrating object. If you touch any object producing sound you can probably feel those vibrations. The higher the pitch of the sound, the faster the rate of vibration of the object. The object's vibration transfers some of its energy of movement to the medium that carries the sound wave or vibrational energy away from the source.

A compression is a zone where the particles of the medium are pushed closer together. It is a zone of higher pressure.

A rarefaction is a zone where the particles of the medium are spread further apart. It is a zone of lower pressure.

The drum is a good example of a device acting as a source of vibrational energy. As shown in figure 2.2, the back and forth vibrations of the drum skin produce air-pressure differences. This produces a vibration effect in the air particles that results in zones of high air pressure (**compression**) and zones of low air pressure (**rarefaction**).

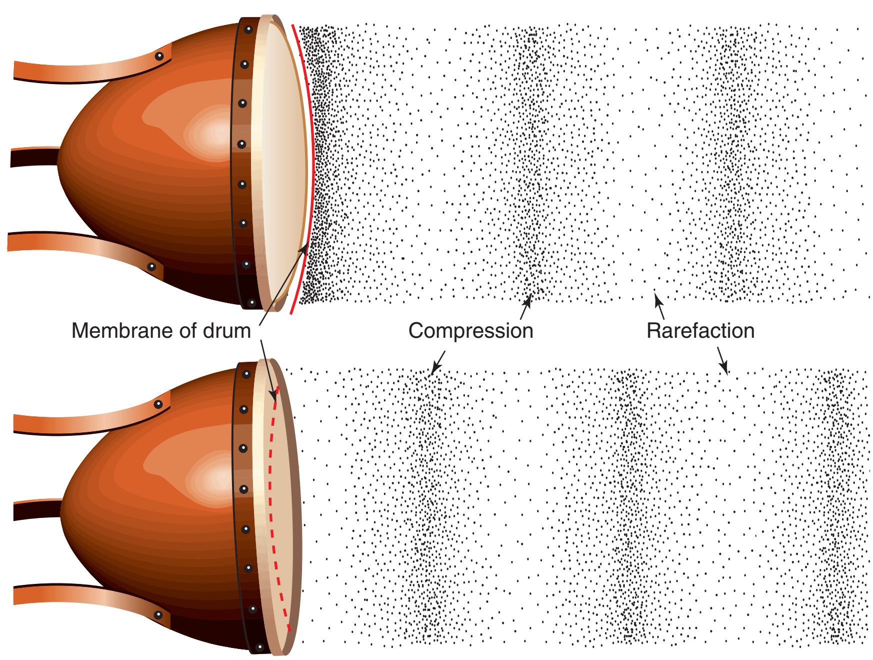


***Figure 2.2** Production of a sound wave in air by the vibrating skin of a drum*

The speed of sound varies in different mediums and within a medium if conditions are changed. The speed of sound is lower in air if the density of the air is lower. For example, on humid days air is less dense, so the speed of sound is slower.

Sound can travel within solids or liquids as well as in gases. You can feel or hear the vibrations you know as sound in a solid simply by touching the solid or placing your ear against it. Consider the following examples.

- Victims of earthquakes who are trapped under rubble are often found because, rather than shout, they tap on a solid material and generate sound waves within it. The sound travels more easily through the overlying rubble within the solid material.
- The approach of a train can be detected in the track long before the sound of the train is heard through the air.
- The song of a whale can be heard hundreds of kilometres away with sensitive microphones designed to pick up those sounds underwater. Whales are thought to use this property of sound to communicate with each other over vast distances.

In all cases where sound waves are transmitted, the sound wave propagates through the material as a vibration of the particles in the form of a pressure wave. As mentioned in chapter 1, that pressure wave consists of compressions and rarefactions of the particles of the medium.

2.2 'SEEING' SOUND WAVES

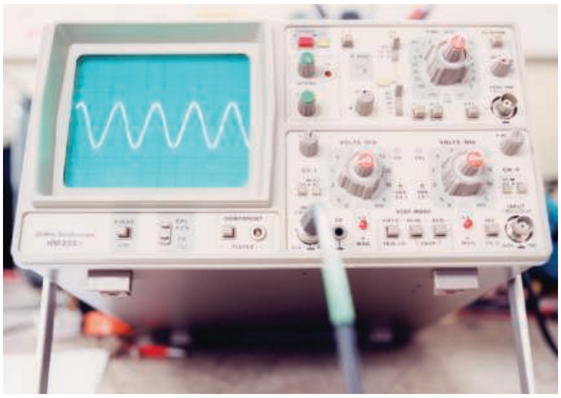
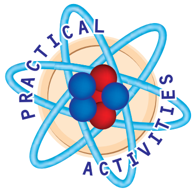


Figure 2.3 *A cathode-ray oscilloscope*

The cathode-ray oscilloscope (CRO) is a device that allows us to view sound waves on a screen (see figure 2.3). The CRO plots or traces the amplitude of the input waveform against time and displays the wave shape on the screen by means of a cathode-ray tube.

Figure 2.4 shows the trace of a sound wave from the screen of a CRO. The areas where the displacement of the wave is above the base line represent zones of compression. The areas where the displacement of the wave is below the base line represent zones of rarefaction. In reality, what happens is that the sound-wave energy is converted into an electrical signal at the microphone. The size of the electrical voltage induced at the microphone is a function of the pressure of the air striking the microphone diaphragm. That pressure differential changes the voltage to a higher or lower value as it passes into the CRO. That voltage input registers on the screen as a trace of a waveform — providing the trace of the sound signal.



2.1

Analysing sound waves from a tuning fork

You may have noticed how the shape of the trace is a sine wave. This is because this wave trace was produced by a constant-frequency source (tuning fork) that produces a pure tone. The wave's traces are hence symmetrical about the base line. Often naturally occurring sounds are not symmetrical because they are not pure sounds of only one tone. The base line represents silence in this case. A trace showing only the base line would indicate no sound at all.

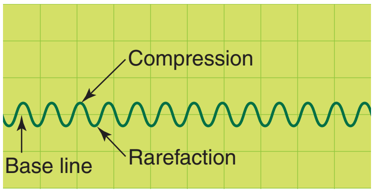


Figure 2.4 *A CRO trace of sound from a tuning fork. Each horizontal grid division represents a unit of time.*

Representing sound waves as transverse waves

Sound waves are often shown in diagrammatic representations as transverse wave traces. You know that these representations are only models because sound waves are not transverse in character. They are longitudinal compression waves.

Earlier in this chapter the CRO trace of a sound wave was described as being above the base line (crest) when the sound wave pulse was a compression and below the base line (trough) when the sound wave pulse was a rarefaction. This is the general situation that occurs when compression waves (such as sound) are represented as transverse waves. The relationship between compressions and crests, and rarefactions and troughs is shown in figure 2.5.

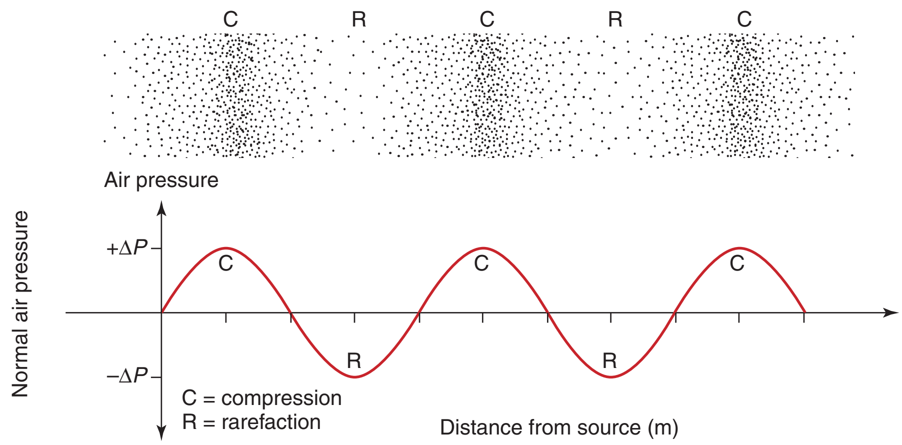


Figure 2.5 *The relationship between compressions and crests and rarefactions and troughs. Notice how the crest of the transverse wave occurs at the centre of the compression where the pressure is at a maximum and the trough of the transverse wave occurs at the centre of the rarefaction where the pressure is at a minimum.*

2.3 AMPLITUDE AND PITCH

Amplitude is the maximum size of the particle displacement from the undisturbed state. In relation to sound waves, the higher the amplitude, the louder the sound.

A sound that is loud — in other words, a high-volume sound — is said to have a high or large **amplitude**. A soft or low-volume sound is said to have a low amplitude. This relationship is often expressed in general conversation in just those terms. We refer to the device that changes volume or loudness on the music system as the amplifier. If a sound wave of a particular frequency, say one generated by a tuning fork, is brought near a microphone connected to a CRO and the tuning fork is made to emit a loud sound by a hard strike, the amplitude of the sound trace on the CRO is much greater than if a soft sound is emitted by the tuning fork (see figure 2.6). This is because the voltage (electrical energy) induced into the microphone is much greater and in proportion with the energy of the loud sound wave. The soft sound wave would produce a much lower amplitude trace. It takes more energy to produce a large amplitude sound of the same frequency.

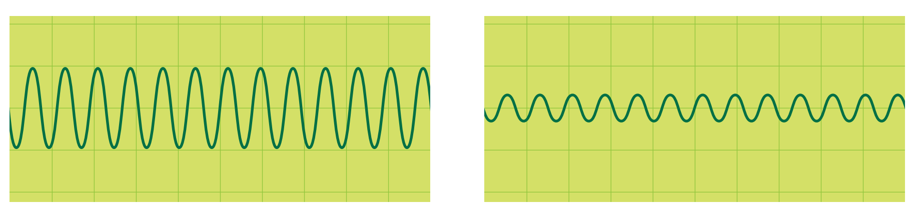


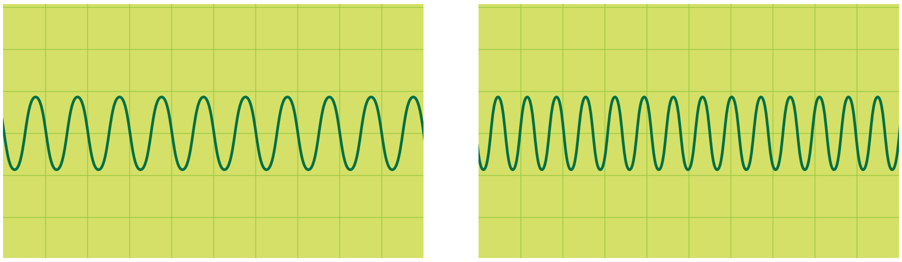
Figure 2.6 *Two CRO traces produced by sounds of the same frequency from a single tuning fork. The trace on the left was from a loud sound, hence the amplitude is quite large, whereas the sound that produced the trace on the right was much quieter and hence has a lower amplitude. The time base represented by horizontal grid divisions on the figures is the same.*

Note that in figure 2.6 the same number of wavelengths are shown for each screen; only the wave height changes.

Pitch is directly related to the frequency of a sound. The higher the frequency of the sound, the more vibrations per second and the higher the pitch. A low-frequency sound is a low-pitched sound.

The **pitch** of a sound is directly related to the frequency of the sound. The higher the frequency of the sound, the more vibrations per second and the higher the pitch. A low-frequency sound is a low-pitched sound. If two different frequency (Hz number) tuning forks, producing sounds of equivalent amplitude, are used to produce CRO traces under identical conditions, the amplitude of the CRO trace waves are equal but the frequency is higher for the higher pitched sound (see figure 2.7).

Figure 2.7 The CRO traces of two sound waves with identical loudness (amplitude) but different frequencies. The figure on the left shows a low-pitch, lower frequency sound-wave trace while the figure on the right shows a high-pitch, higher frequency sound-wave trace. The time base represented by horizontal grid divisions on the figures is the same.



2.4 ECHOES: REFLECTIONS OF SOUND

An **echo** is a repeated sound created by the reflection of sound waves from a surface.

You are familiar with the sound of an **echo**. You are probably aware that an echo is sound reflecting from a surface and bouncing back at you. Echoes are heard when you are some distance from the surface that reflects the sound, as shown in figure 2.8.

When an echo bounces back from a solid surface, such as a cliff face or a brick wall, you don't hear the full sound, but you do hear the last part of the original sound. If you are a significant distance from the wall, you will hear more of the original sound bounce back. If you are close to the reflecting surface, you probably won't detect an echo. It does still occur, but the original sound drowns it out. There needs to be a time difference between the reflected sound and the original sound so that you can hear the echo. The size of that time difference is a minimum 0.1 seconds. Because sound travels around 340 m s^{-1} in air, both you and the sound source must be at least 17 metres from the surface reflecting the sound for you to hear the echo. At this distance, the sound wave takes 0.05 seconds to reach the reflecting surface from the sound source and 0.05 seconds to bounce back.

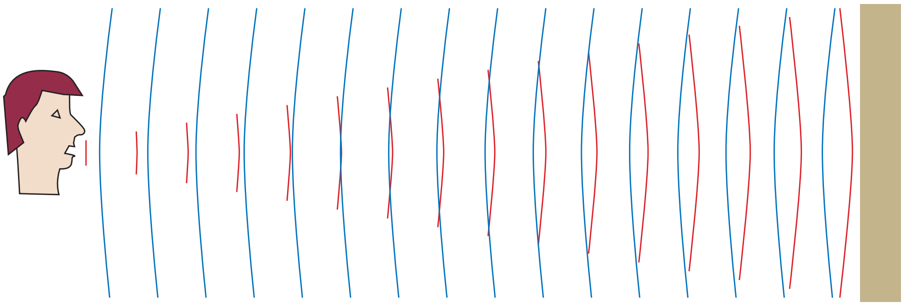


Figure 2.8 An echo bouncing back from a wall. Note that the wave bounced back is out of phase with the wave hitting the wall. This phase change upon reflection is a characteristic of all waves.

Echoes are used by sonic rangers to determine the distance to objects. In the water, sonar and depth finders on boats are used to determine the distance to objects underwater or to the floor of the ocean (see figure 2.9). Sonic rangers are also used by industry in sonic level controllers to tell how full storage tanks are. In most of these applications it is desirable to use short-wavelength, high pitch sound waves. These ultrasonic or very high-frequency sound waves are emitted from a source and bounce back from objects. After bouncing back, they are detected by pressure-sensitive detectors. The time they take to return to their source can be determined accurately. Knowing the time for the reflected wave or echo to be received allows the distance to the object to be calculated. When calculating the distance, the speed of sound in a medium is assumed to be constant (even though slight fluctuations in the speed of sound do occur as the density changes).

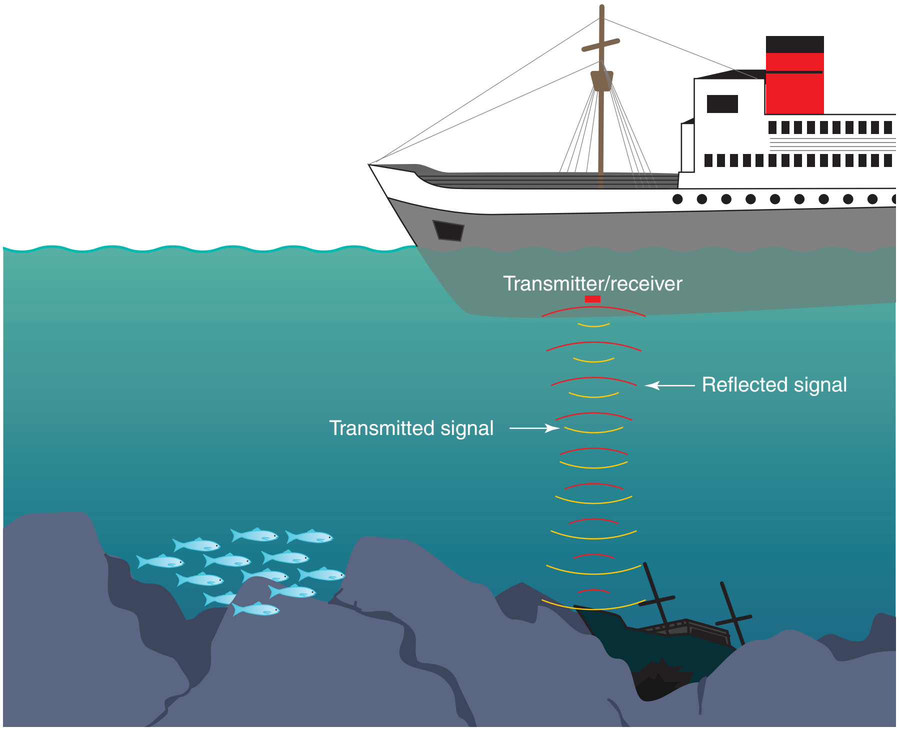


Figure 2.9 *Depth finders on boats use sound echoes to measure the depth to the ocean floor or to other objects or fish.*

Because there is a time difference between reflections of the same pulse if the reflecting surface is irregular in shape, it is possible to use ultra-high-frequency sound waves to 'see' objects. The reflections from multiple surfaces are processed by a computer to generate an image of the object's surface. This technology is used extensively in medicine to perform non-invasive examinations to determine soft tissue injuries, diseased organs or to check unborn children for abnormalities (see figure 2.10).

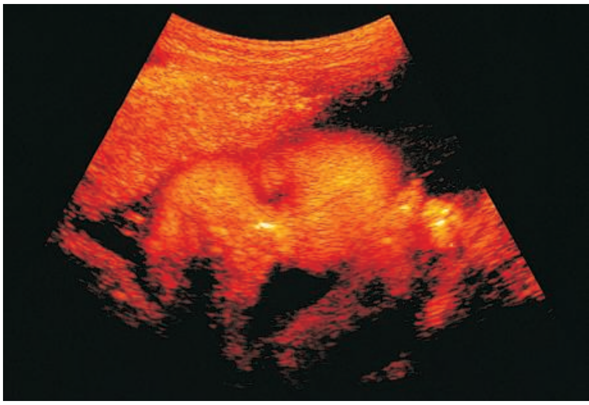
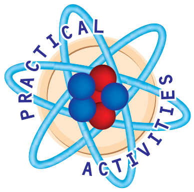


Figure 2.10 *An ultrasound image of a baby in the womb*

2.5 SOUND AND THE PRINCIPLE OF SUPERPOSITION



2.2

Observing wave interference

When waves meet they interact as they pass through each other, reinforcing or cancelling at different points. This is called interference.

Sound waves from separate sources **interfere** with each other and their amplitudes can be added. When this occurs it is possible to produce a sound of higher or lower amplitude depending on whether the sound waves are in or out of phase. If the amplitude of the crest of one wave is precisely equal to the amplitude of the trough of another wave, and the second wave is out of phase by 180° from the first, then annulment or complete loss of amplitude in the resulting sound wave produced by their addition can occur. This means it is possible to add sounds together to produce no sound. If the waves are in phase, and troughs coincide with troughs and crests coincide with crests, the amplitude of the wave produced will be increased. This means that the sound will be louder.

Adding waves — superposition

Superposition is the adding of two or more waves.

The addition of waves is called **superposition**. Given two (or more) waves we can determine the resultant wave graphically. To do this we must apply the superposition principle which states that if two or more waves of the same type pass through the same medium at the same time, then the amplitudes of the waves add together. The principle involves adding individual displacements at various points in a systematic way.

SAMPLE PROBLEM 2.1

Adding waves

Add the two waves graphed in figure 2.11.

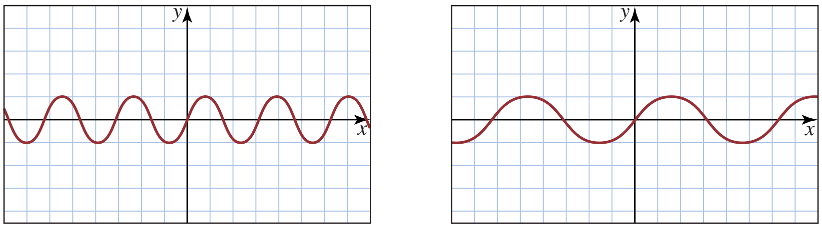


Figure 2.11

SOLUTION

The resulting graph is shown in figure 2.12.

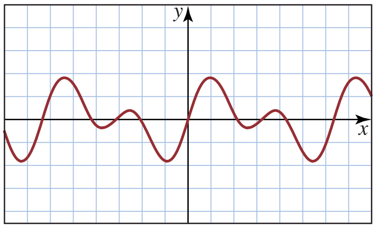


Figure 2.12

SAMPLE PROBLEM 2.2

Adding waves

Add the two waves graphed in figure 2.13.

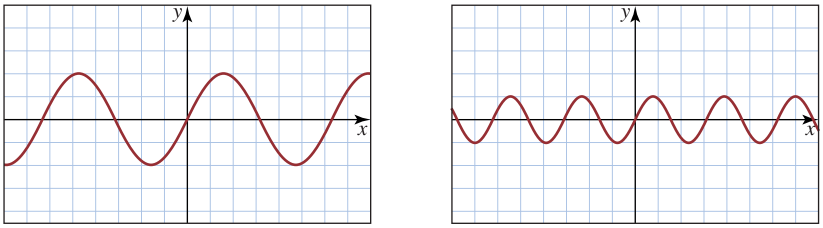


Figure 2.13

SOLUTION

The resulting graph is shown in figure 2.14.

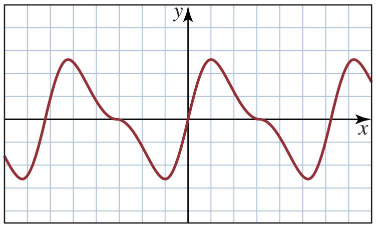


Figure 2.14

As you can see, the addition of rather simple wave shapes can form a complex wave.

SAMPLE PROBLEM 2.3

Adding waves

Add the two waves graphed in figure 2.15.

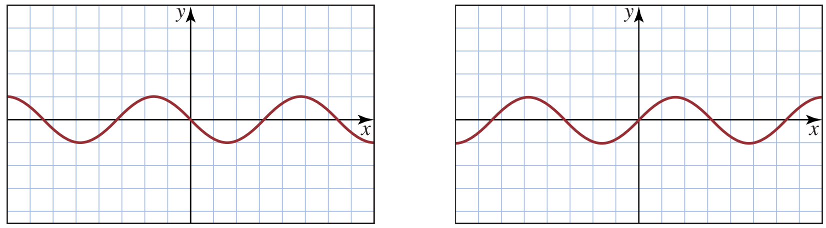


Figure 2.15

SOLUTION The resulting graph is shown in figure 2.16.

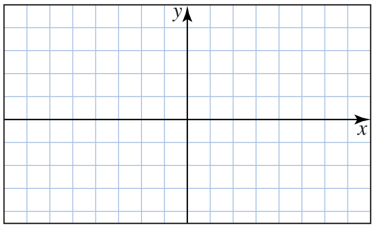


Figure 2.16

This wave shows annulment of the waves. The two waves added were out of phase by 180°.

If two out-of-phase sound waves interfere, the amplitude of the resulting sound wave will be less than either of the original waves. If two in-phase sound waves interfere, the resultant sound wave will have a greater amplitude than either of the original waves. These phenomena are often apparent in everyday life. Some examples are outlined here:

- The development of technology has enabled the production of sound waves that are out of phase by 180°. This has been used to reduce the sound emitted by heavy machinery in factories.
- This technology is used in some expensive headphone sets to eliminate the effects of external sounds, enabling the wearer to hear recordings cleanly despite the noise of surroundings.
- The addition of sounds and their echoes in enclosed stairwells of multi-storey buildings often means that the stairwell contains dead spots, or spots where the sound amplitude is much greater. You may have experienced this when walking up or down stairs while also listening to a sound emitted by a stationary source within the stairwell. The apparent amplitude of the sound varies from loud to soft as you climb up or down the stairs, even though the actual amplitude of the emitted sound from the source does not vary.

Interference of waves can also produce waves that are of a different frequency to the waves that produced them. If two waves with different frequencies are added, the result can be a wave that is somewhat more complex in character (see figures 2.12 and 2.14, where the waves are more complex than, and have a different frequency to, the original source waves).

Beats — a special case of superposition

When two sources of sound of the same amplitude but slightly different frequency are heard together, there will be a rhythmic change to the volume of the sound. When the two sound waves are in phase, the amplitude of the resulting sound wave is the sum of the amplitudes of the two waves, and results in a loud sound. As the waves drift out of phase, the resultant amplitude will become smaller, eventually reaching zero before increasing again as the waves drift back into phase. The term '**beats**' is used to describe the variation in the loudness of the sound.

The term beats refers to the change in volume of a sound that occurs when two sounds of slightly different frequencies occur together.

For example, beats may occur when members of an orchestra are warming up for a performance and are tuning their instruments. As the tuning of the instruments becomes closer, the beat frequency decreases until it disappears. The beat frequency is determined by the difference in the frequency of the notes played by the different instruments when slightly out of tune.

It is possible to see beats on an oscilloscope as a rhythmic pulsing of wave amplitudes. The waveforms are produced by the superposition of sound waves of similar frequencies from multiple sources.

PHYSICS FACT

PRACTICAL ACTIVITIES

2.3

Analysing sound waves from musical instruments

Timbre: Combining pure tones

The sound produced by a tuning fork is a pure tone. The CRO trace of such a sound is a sine wave, as shown in figure 2.17. Most sounds are not pure tones but are made up of a number of pure tones that have been superimposed in a particular way to produce a sound with a characteristic **timbre** (see figure 2.18).

Timbre is the quality of a sound that depends on the way in which a number of different pure sounds have combined.

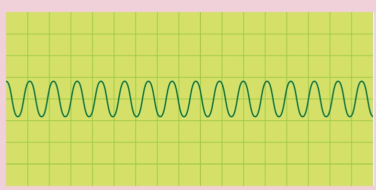


Figure 2.17 *A pure tone produced by a tuning fork. The CRO trace is a sine wave shape.*

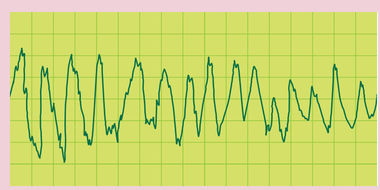


Figure 2.18 *A complex sound produced by a person singing. Note that in any CRO trace figure, the trace represents a very small 'grab' of time, much like a photograph.*

Although the shapes of the waves for the figures above are different, the frequencies are approximately the same. The difference is the timbre, or complexity of the note. This is borne out in the difference in shape.

You are probably aware that different musical instruments playing in an orchestra can play the same musical note. However, while the sounds are of the same frequency, they do not appear to be the same. This is because the sounds produced have their own particular timbre. If viewed as a CRO trace, these common notes from the different instruments produce a differently shaped wave trace even though the frequencies are common. Figure 2.19 shows the wave traces for the same note played by a number of different instruments into a microphone.

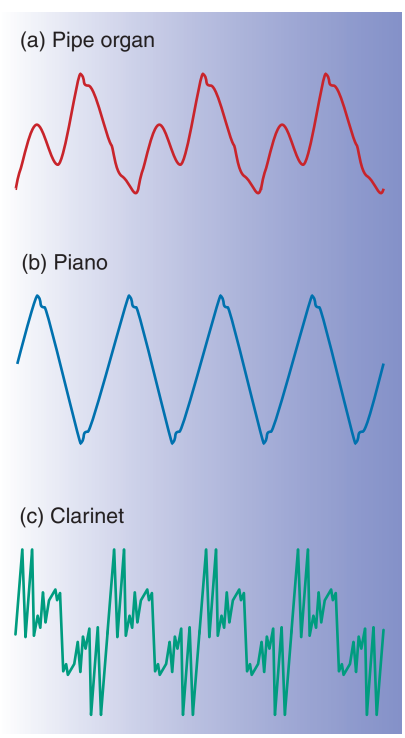


Figure 2.19 *The CRO traces at the same settings of the same musical note played with approximately the same volume by a variety of different instruments: (a) pipe organ, (b) piano and (c) clarinet. These CRO traces are more complex than the sounds produced by tuning forks and often lack symmetry about the base line.*

SUMMARY

- Sound waves are vibrations of particles in a medium.
- Compressions relate to the crests of a transverse wave and rarefactions relate to the troughs of a transverse wave.
- The pitch of a sound wave increases as the frequency of the sound wave increases.
- The amplitude of a sound wave increases as the volume of sound grows louder.
- An echo is a reflection of a sound wave.
- Waves can interfere when they come into contact. That can result in the amplitude of the waves increasing if the waves are in phase or decreasing if the waves are out of phase. Addition of waves is called superposition.
- Sound waves can be studied with a cathode-ray oscilloscope (CRO) or cathode-ray oscilloscope simulator computer program.
- Different musical instruments produce sound waves that produce different shaped traces on a CRO.
- Tuning forks produce pure notes that result in a sine-wave trace on a CRO.
- The notes from tuning forks can be added to produce more complex sounds and wave traces on a CRO. This is an example of superposition of sound waves.

QUESTIONS

1. If sound travels in air at 330 m s^{-1} and a tuning fork producing sound travelling at that speed vibrates with a frequency of 256 Hz, calculate the wavelength of the sound wave produced by the tuning fork.
2. If two sound waves are travelling in the same medium but one sound wave has a frequency twice as high as the other, compare their wavelengths.
3. Much publicity has surrounded the development of sound-eliminator technology whereby a machine or headphone technology that generates a noise is used to eliminate the loud noise of a machine or background sounds. Explain how this is possible.
4. Sound waves leaving one medium enter another medium where their speed is much greater. In both media, the frequency of the sound wave remains constant. Explain what is different between the sound wave in each medium.
5. Describe the principle of superposition.
6. Explain why high-frequency sound waves are preferred for tasks such as echo location rather than low-frequency sound waves.
7. A baritone and a soprano were asked to sing the word 'one' into a microphone and the signal produced was fed into a CRO. Describe how you would assign each of these different traces to the correct person.
8. Calculate how many times per second a tuning fork vibrates if it is producing a sound wave in helium gas with a frequency of 384 Hz.
9. A CRO trace produced by a sound wave is shown in figure 2.20. Identify on the trace where the air particles are:
   (a) undergoing rarefaction
   (b) undergoing compression
   (c) in neither compression nor rarefaction.

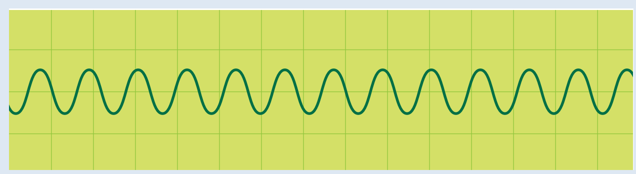


Figure 2.20 ***A CRO trace***

10. Figure 2.21 shows an experiment often done by students learning about sound wave properties.

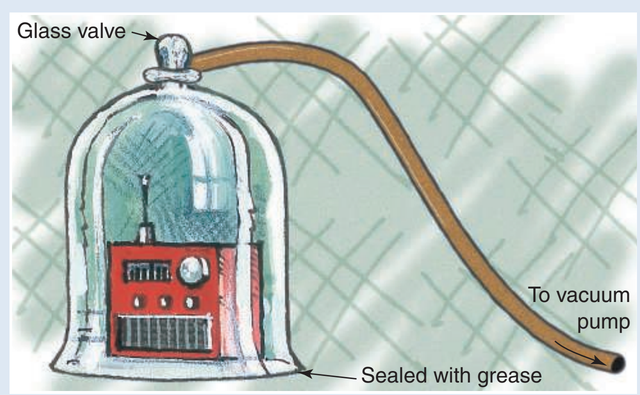


Figure 2.21 ***An experiment used to study sound waves in a vacuum***

The radio is switched on in a large sealed bell jar attached to a vacuum pump. As the air is pumped from the bell jar, the sound of the radio becomes progressively softer and finally disappears. The radio continues to work as normal throughout this experiment. Explain this phenomenon using your knowledge of the properties of sound.

CHAPTER REVIEW

11. Present diagrammatically (on graph paper) the following two transverse waves (that are initially in phase) and add the waves to produce a resultant wave.
Wave 1: wavelength 2 cm, amplitude 1 cm
Wave 2: wavelength 4 cm, amplitude 2 cm
12. Present, as diagrams on graph paper, the following two transverse waves (that are initially out of phase) and add the waves to produce a resultant wave.
Wave 1: wavelength 2 cm, amplitude 1 cm
Wave 2: wavelength 4 cm, amplitude 2 cm
13. Two special tuning forks are used in an experiment with a CRO to determine the trace pattern of a sound wave. One sound wave has a frequency of 250 Hz, the other 500 Hz. The screen of the CRO can capture 0.05 s of the sound. Draw labelled figures to show each of these screen traces.
14. An audio oscillator is a device that can produce sound over a wide range of frequencies and amplitudes. Every frequency produces a sine wave trace on the CRO. A human voice box can also produce sounds over a wide range of frequencies and amplitudes yet the CRO trace is rarely a sine wave. Explain the difference between the two traces.
15. Explain why you cannot hear an echo from a very close wall yet a microphone attached to a CRO can detect an echo as a shadow trace.
16. Describe the production of dead spots, where sound cannot be heard in a room or stairwell despite the fact that a constant noise is being generated.
17. If two waves were superposed and the trace of the resultant wave was a trace along the baseline, what has happened to the sound?
18. Predict whether sound would travel faster in air or water based on the nature of the wave and the medium. Justify your prediction.
19. Explain why high frequency sound waves ‘see’ better than low frequency sound waves when using sonar.
20. Identify which type of sound wave would carry the most energy: a low amplitude sound with a frequency of 256 Hz or a high amplitude sound with a frequency of 512 Hz.

2.1 ANALYSING SOUND WAVES FROM A TUNING FORK

Aim

To observe and collect sound traces from a CRO

Apparatus

at least two tuning forks of different frequency
access to a CRO or a CRO simulation program for the computer
a microphone to convert the sound wave into an electrical signal

Theory

The traces from a CRO can provide you with a snapshot of a number of different sound waves. The waves are a small time-grab of a much larger train of sound waves. These short interval grabs can show you some of the features of a sound.

Method

1. Connect the microphone to the input of the CRO or the microphone input on the computer if using a CRO simulation program.

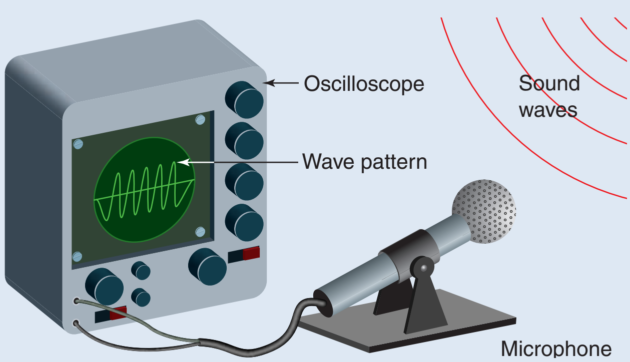


Figure 2.22 ***A microphone attached to a CRO***

2. Tune and adjust the CRO so that when a single tuning fork is brought near to the microphone, a sine-wave trace is produced. Observe what happens to the amplitude of the wave as the tuning fork loses its vibrational energy and the sound becomes softer.
3. Check out all of the traces of all tuning forks you have. When doing this, keep the same CRO settings to make comparison easier. Note the frequencies and shape of the waves produced. If you are using a CRO simulation computer program, you should be able to freeze the CRO traces, save them and print them out.
4. Try striking two different frequency tuning forks and having the microphone collect the sound from both tuning forks. You will notice the shape of the CRO trace wave becomes more complex.
5. Try adding a third sound from another tuning fork to the input into the CRO. Observe the increasingly complex CRO trace.

2.2 OBSERVING WAVE INTERFERENCE

Aim

To hear sound waves interfering with each other

Apparatus

tuning fork

Theory

Each of the vibrating tuning fork prongs acts as a coherent source of sound because it has the same frequency, amplitude and phase in relation to the other when producing a sound wave in air. Hence, there are two sound waves generated by the tuning fork prongs. Each one radiates from a slightly different position. As a compression is produced between the prongs, a rarefaction is produced outside each of the prongs and vice versa. The sound waves propagate outward from each tuning fork prong but on some paths they overlap. This is either because there is a full wavelength difference in the travel path length or because in some directions they meet at a point one half wavelength out of phase. In these directions where the sound waves are exactly one half wavelength out of phase (compression meets rarefaction) the sound waves will add. If the amplitudes are the same, one sound wave's compression is annulled by the rarefaction from the other. This produces a sound minimum.

The sound waves can add to form a maximum if the path difference is equal to a whole number of wavelengths. The result is a higher amplitude sound.

Method

1. Strike a tuning fork so that it produces a note.
2. Hold the tuning fork to your ear and rotate the tuning fork about its long axis.
3. Listen carefully to the sound you hear. Note when the sound waves appear to increase in amplitude and decrease in amplitude.

2.3 ANALYSING SOUND WAVES FROM MUSICAL INSTRUMENTS

Aim

To observe the sound from musical instruments on a CRO or CRO computer simulation program

Apparatus

access to a CRO or a CRO simulation computer program
microphone
variety of musical instruments

Theory

The notes produced by a musical instruments have a characteristic timbre. An examination of the CRO traces of the same note played by a number of different musical instruments will highlight the differences in the nature of the sound waves.

Method

1. Connect the microphone to the input of the CRO or the microphone input on the computer if using a CRO simulation program.
2. Tune and adjust the CRO so that when a single tuning fork is brought near to the microphone a sine-wave trace is produced.
3. Using a variety of musical instruments, play the same note into the microphone attached to the CRO and observe the differently shaped wave patterns that are produced. If you are using a CRO simulation computer program, you may be able to freeze and save the CRO traces and then print them out.

Analysis

Compare the shapes of the CRO traces for each instrument. Are there any similarities? Which instruments are most similar?

CHAPTER 3

ELECTROMAGNETIC WAVES AND COMMUNICATION

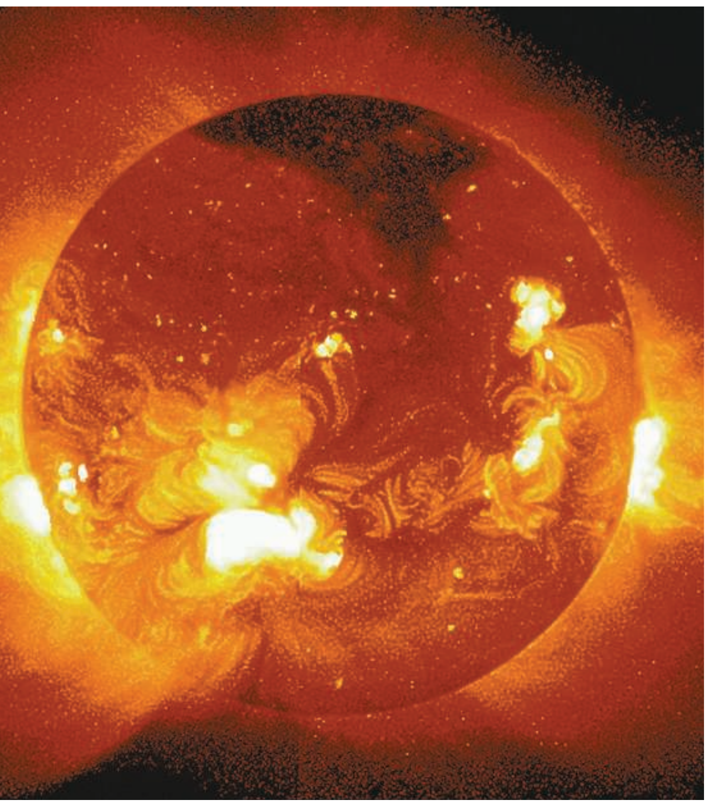


Figure 3.1 *The Sun is the main source of electromagnetic radiation reaching the outer atmosphere of the Earth. The Sun emits the full spectrum of electromagnetic radiation. Only the Earth's atmosphere prevents the catastrophic destruction of life as we know it by absorbing harmful electromagnetic waves.*

Remember

Before beginning this chapter, you should be able to:

- recall some types of radiation that make up the electromagnetic spectrum and identify some uses for them.

Key content

At the end of this chapter you should be able to:

- identify electromagnetic waves in terms of their speed and their lack of requirement of a medium for transmission
- identify how the atmosphere acts as a filter for electromagnetic wavebands, especially UV radiation, X-rays and gamma rays
- identify ways to detect particular wavebands from the electromagnetic spectrum
- explain that the relationship between a change in intensity of light and other examples of electromagnetic radiation and a change in distance from the source obeys the inverse square law $I \propto \frac{1}{d^2}$
- outline how information can be transmitted by modulating the amplitude or frequency of visible light, microwaves and radio waves
- discuss the limitations of electromagnetic waves when used for communications
- identify electromagnetic waves involved in the transfer of energy when using radar.

3.1 THE WAVES OF THE ELECTROMAGNETIC SPECTRUM

The **electromagnetic spectrum** is the full range of wavelengths of all electromagnetic waves.

The Sun, Earth and most other bodies in our universe radiate electromagnetic energy. Electromagnetic waves do not require a medium in order to travel from place to place. In fact, they travel most efficiently in a vacuum, such as space. Electromagnetic waves travel at the speed of light, can be reflected and refracted, and can carry information as codes.

The **electromagnetic spectrum** is a continuum of electromagnetic waves with artificial divisions based on the frequency and wavelengths of the waves. There is no distinct point at which the frequency changes and no special change in properties at particular wave boundaries. The concept of a continuum can be illustrated by looking at a rainbow. As figure 3.2a shows, boundaries between the individual colours are not clearly defined yet we can see a transition from one colour to the next.

Figure 3.2b shows the wave types that make up the electromagnetic spectrum and the wavelengths and frequencies of the different wave types.

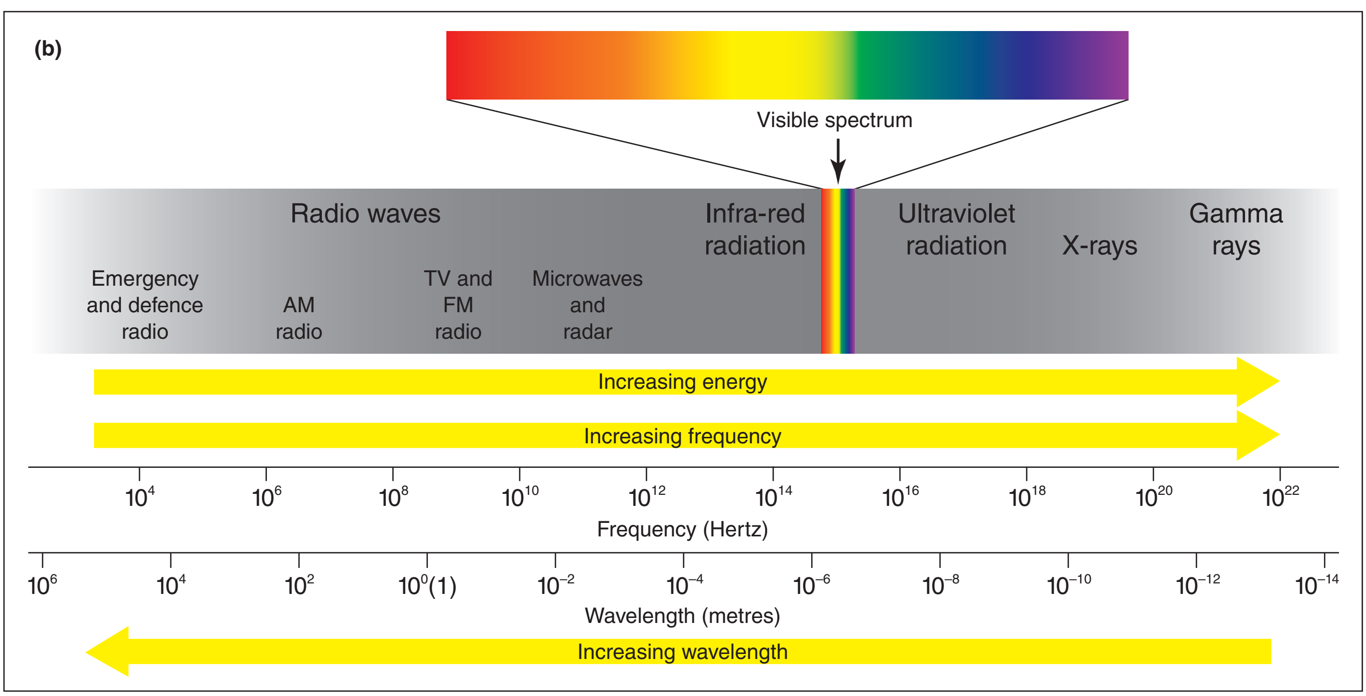


Figure 3.2 *(a) The continuum of the spectrum of visible light is apparent in the rainbow. (b) The electromagnetic spectrum*

Even before the speed of light was accurately measured, or the full electromagnetic spectrum was known, James Clerk Maxwell (1831–1879) predicted that the speed of all electromagnetic waves was the same.

Electromagnetic waves have special properties that are outlined below:

- All electromagnetic energy passes through the vacuum of space at the common speed of light (300 million m s^{-1} or 3×10^8 m s^{-1}).
- The waves are produced by oscillating, perpendicular electric and magnetic fields, hence the term electromagnetic (see figure 1.9, page 8).
- The waves are considered to be self-propagating; that is, the electric field produces a perpendicular magnetic field that in turn induces an electric field and so on. This property enables electromagnetic waves to travel the immense distances of the universe for billions of years away from the source of the radiation.
- As a general rule the electromagnetic waves are represented in the form of sinusoidal (sine-shaped) or complex transverse waves. These complex electromagnetic waves are like sound; that is, they are made up of multiple simple waves superposed.
- They are capable of creating an electrical response in the medium that they impact.
- They have frequencies related directly to the vibrational frequency of their source particle.
- They can be reflected by certain materials and can be refracted or bent when passing from a medium of one density to a medium of a different density.

All of these properties are used in the communication technologies that use electromagnetic waves to carry their signals.

Electromagnetic waves used in communication technologies

The parts of the electromagnetic spectrum that we use for communication include radio waves, microwaves, infra-red, visible light and ultraviolet.

Radio waves

As well as their role in communication technologies, radio waves of a few centimetres in wavelength are used in producing radar maps. If radio waves are transmitted from a satellite or plane, they can bounce off the surface of the Earth to produce a reflection which, when detected by suitable equipment, can create a picture of what lies below.

Radio waves have wavelengths ranging from 10 cm to 1000 m and are assigned to the AM, FM, VHF or UHF category based on their frequency. Radio waves are used to transmit television, FM and AM radio, radar and some mobile telephone signals. Tuning your television or radio to a particular station means that you are tuning to a particular frequency range or band that matches its signal, for example, Triple M at 104.9 MHz.

PHYSICS FACT

Electromagnetic radiation can be generated by accelerating charged subatomic particles (for example, electrons) to produce an effect similar to that of a vibration in a source of mechanical waves. The radio-wave signal produced by a radio transmitter is the result of electrons being forced to accelerate up and down the length of the aerial. The rate of vibration of the charged particle determines the frequency of the electromagnetic radiation produced by the particle. The electron vibrations in a radio aerial producing a signal in the FM radio band (with frequencies in the MHz (megahertz) range) are much more rapid than those in an aerial producing an AM radio-signal broadcast (in the kHz (kilohertz) radio band). The greater the acceleration a charged particle experiences or the more rapidly it is forced to change direction in the device acting as an aerial, the shorter the wavelength of the electromagnetic radiation produced.

Figure 3.3 *Some applications of microwave technology (a) Mobile phone signals (b) A microwave oven*

Microwaves

Microwaves have wavelengths between one millimetre and 30 centimetres. One application of microwaves is the microwave oven that emits waves tuned to a frequency of 2450 MHz. The water in food absorbs the energy of the microwaves and becomes warmer.

Microwaves are emitted from the Earth because the Earth is a relatively hot body. They can also be emitted from objects such as cars and planes, and from the atmosphere. These microwaves can be detected and interpreted to give information, such as the temperature of the object that emitted them.

Microwave links are used to transmit mobile phone signals on frequencies of around 900 MHz. Transmission can be across distances of up to 100 km, but there must be a direct 'line of sight' from the microwave transmitter to the receiver dish because microwaves travel in straight lines.

Infra-red radiation

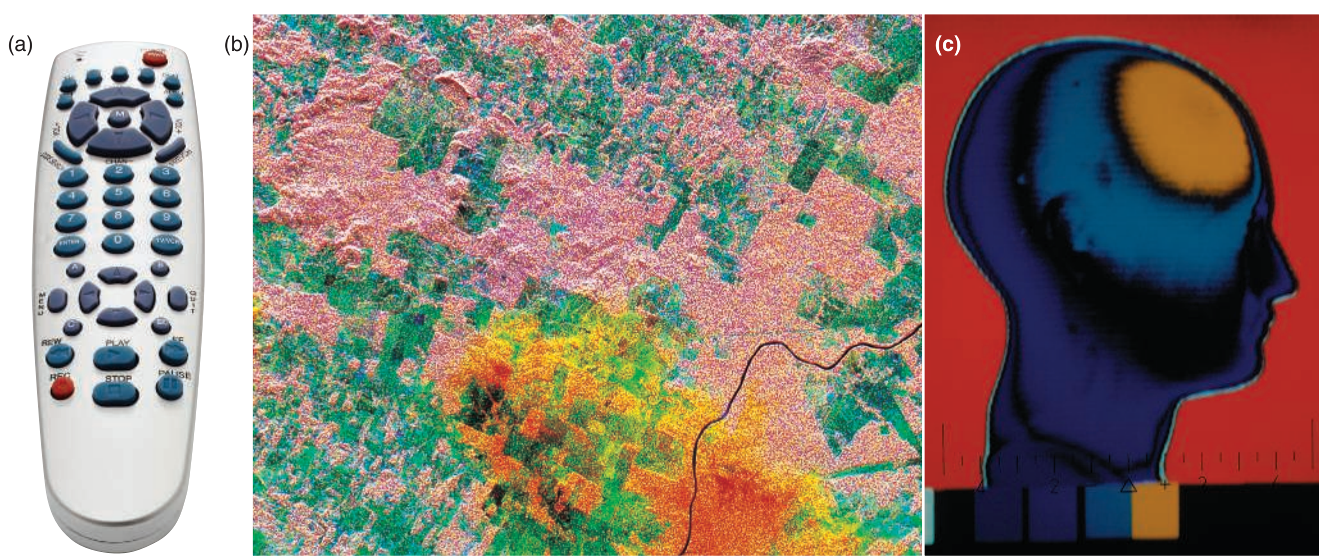
Infra-red waves have wavelengths of around 700 nm to 1 mm. Infra-red radiation levels can be measured using electronic detectors.

1 nanometre (**nm**) *=* 10^{-9} **m**

Many electronic remote controls, such as those for garage doors or the television, and remote-control wireless connections to computers, use infra-red radiation to transfer the signal from the control device to the consumer item. Data can be broadcast over relatively short distances from computer to computer by infra-red signal without the need for wiring. In the telecommunications industry, infra-red lasers send information down optical fibres.

Other applications of infra-red radiation include medical treatments for soft tissue injury and finding heat leaks from houses. Switchboards in large buildings are photographed using infra-red cameras to check for hot spots that might indicate dangerous electrical faults or that electrical circuit overloading is occurring. These problems cause the wiring to carry too much electrical current and thus heat up so that the wire gives off infra-red radiation. Satellite infra-red images can give farmers information on the health of crops. Aeroplanes that carry infra-red detectors can identify forest-fire hotspots even when hidden by a curtain of smoke.

Figure 3.4 *Some applications of infra-red radiation (a) The television remote control (b) Remote sensing of vegetation patterns (c) An infra-red photograph of a person's head*



Visible light

Visible light has a wavelength of 400–700 nm. Applications include: fibre-optic telecommunications, remote sensing of vegetation patterns from satellite and aeroplane surveys, and identification of different objects by their visible colours.

Visible light is certainly the most important electromagnetic wave type used in communication. The use of reflection and visible signalling, such as smoke signals, was almost certainly the first long-distance method of communication.

Ultraviolet radiation

Some scientists believe that the crew of aeroplanes have an increased risk of cancer and that male crew have lower sperm counts because of their higher exposure to high energy electromagnetic rays at high altitudes.

When participating in sporting activities such as skiing or mountaineering it is important to protect the skin and eyes. The increased risk of sunburn, even in the cold, is due to the greater level of penetration of the UV at higher altitudes.

Ultraviolet (UV) radiation has wavelengths of 10–400 nm. A small dose of this radiation is beneficial to humans because it encourages the production of vitamin D, which is essential for strong bones. Larger doses of UV radiation can lead to cell and tissue damage, possibly causing skin cancer or eye cataracts that severely affect vision. Most harmful UV radiation is absorbed by the Earth's atmosphere before it can reach the Earth (see figure 3.5. page 33).

The use of UV to cure plastics is becoming increasingly common. Liquid plastic components are designed to contain photo-initiator chemicals that start the polymerisation reaction in plastics. This reaction does not produce gases as a by-product, so it is useful where maintaining transparency is important, such as repairing cracks in windshields.

Other applications of UV radiation include its use in making astronomical observations and for sterilising hospital equipment and hairdressers' combs. High-output UV lamps can be used to sterilise water in water purification systems, and in fish tanks a high-UV-emitting lamp sterilises the upper few centimetres of water, killing off unwanted algae. Because of the limited penetration of the UV into the water, the fish are unharmed by this technique because they spend little of their time within the high-UV-penetration zone.

Other waves of the electromagnetic spectrum

X-rays

X-rays have wavelengths of 0.01–10 nm. Applications include the standard medical examination of hard or dense parts of the body without the need for surgery. In engineering, X-ray devices are used to inspect welds for cracks. Because X-ray wavelengths are of the same order of magnitude as the size of atoms, they can be used in X-ray crystallography to see the structure of crystalline materials. It is important to study X-rays from space so we can predict space weather. X-rays are also used in the manufacture of electronic microchips and in the efficient production of many biomolecular materials. X-rays can be detected with photographic film.

Gamma rays

Gamma rays have wavelengths of less than 0.01 nm. They can have medical applications, such as in the treatment of some cancers in radiation therapy, and are used in astronomical investigations of high-energy events in space. Such events include the study of remnant signals from past supernovae. Gamma rays can be detected with a Geiger counter.

3.2 ATMOSPHERIC FILTERING OF ELECTROMAGNETIC WAVES

The Earth's atmosphere and ionosphere absorb most of the incoming electromagnetic radiation from space except for visible light and some high-frequency radio waves in the microwave region. The other types of radiation would generally be harmful to us. Too much UV radiation, for example, would cause cancer and dangerous mutations. Too much penetration by X-rays or gamma rays would quickly kill us.

When using Earth-bound instruments to study the universe (such as optical telescopes and radio telescopes) astronomers make extensive use of radiation in the visible and radio range. To take advantage of the entire electromagnetic spectrum for observations, astronomers are increasingly using space satellites and probes to detect the bands of the electromagnetic spectrum that do not reach Earth. Light telescopes such as the Hubble Space Telescope are providing brilliant, high-resolution optical images from beyond the atmosphere.

The seasonal variation in the concentration of UV-absorbing ozone in the atmosphere affects UV penetration.

PHYSICS IN FOCUS

Chandra X-ray telescope

The Chandra X-ray telescope is designed to study the X-rays produced when gas is heated to millions of degrees by violent and extreme conditions in flaring stars, exploding stars, black holes and vast clouds of hot gas in galaxy clusters. Images from the Chandra telescope show 50 times more detail than any previous X-ray telescope based on the Earth's surface because the telescope sees 50 times more X-ray information. The Chandra telescope was placed in an orbit one-third of the distance to the Moon. This was to ensure minimum absorption of the X-rays by the atmosphere and to escape the influence of the Earth's magnetic field that affects the X-ray-detecting instruments.

Microwave penetration decreases as the amount of water vapour in the air increases. This produces variation on an almost daily basis for this important communication wave-type that is used to carry the signal from mobile telephones.

The amount of any particular frequency of electromagnetic waves that can penetrate through the atmosphere to the surface of the Earth varies seasonally and with space and atmospheric weather conditions. Because of this variation, every frequency of the spectrum does reach the Earth's surface to some degree. Figure 3.5 shows the approximate zone where most of any particular electromagnetic wave frequency is absorbed in its passage toward the Earth's surface.

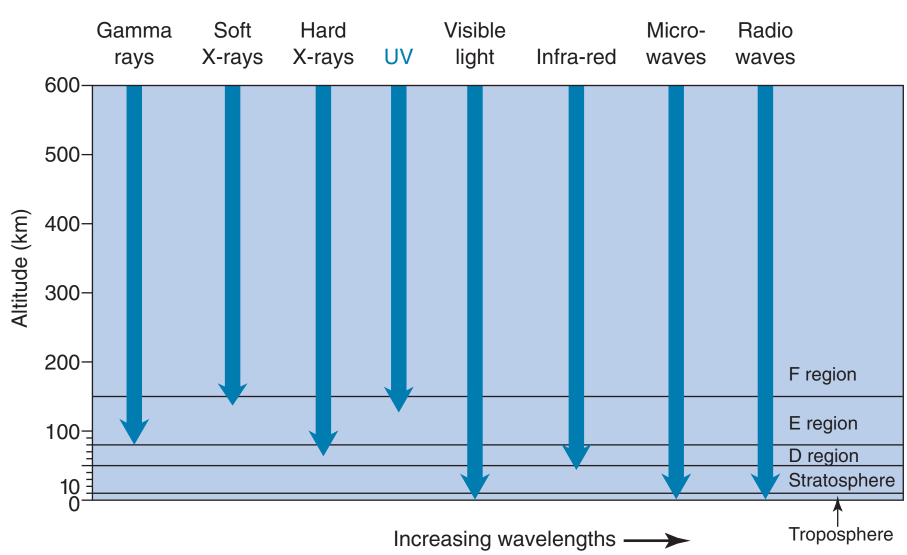


Figure 3.5 *The layers of the atmosphere and ionosphere and the relative penetration of electromagnetic waves through these atmospheric layers. Where the arrow stops indicates where most of that type of radiation ceases to penetrate.*

Figure 3.6 *This antenna receives satellite radio transmissions.*

A negligible amount of some electromagnetic waves, such as X-rays, penetrates the atmosphere to reach the Earth's surface. Astronomers who use ground-based detecting equipment need large collecting and focusing devices to collect the weakly penetrating electromagnetic frequencies.

Satellite dishes receiving signals from space satellites also have large surface areas (see figure 3.6).

The ionosphere as a filter

A layer of gas called the atmosphere surrounds the Earth. At high altitude, roughly between 50 km and 500 km above Earth, part of that gas is ionised; that is, by losing or gaining electrons the atoms and molecules of gas have become charged. These ions form the layer of the atmosphere referred to as the ionosphere.

The ionosphere can be divided into three layers (D, E and F) based on the type of electromagnetic radiation absorbed in each layer (see figure 3.5 on page 33).

The regions of the ionosphere have the following characteristics:

- The D region extends about 50–80 km above the Earth's surface. Hard X-ray radiation with short wavelengths and high frequencies is absorbed in the D region.
- The E region extends from around 80 km above the Earth's surface and peaks at an altitude around 105 km before finally being replaced by the F region at an altitude of 145 km. Soft X-rays (those with longer wavelengths) are absorbed in the E region.
- The F region extends from around 145 km and continues to about 300 km above the Earth's surface although it may reach a height of up to about 600 km at night. Extreme ultraviolet radiation with short wavelengths such as UVc (see page 36) is absorbed in the F region.

The exact height of the regions actually varies during the day with the fluctuations in received solar radiation. The D region almost disappears at night but the other two regions rise and become less strongly ionised as the amount of incoming solar radiation drops.

Some changes that occur in the ionosphere can affect communications on Earth. Generally, the upper layers of the ionosphere reflect radio waves below some critical frequency (about 30 MHz) while the lower layer acts more as an absorber of radio waves. However, some solar activities, such as solar flares, result in dramatic bursts of energy and increase the electromagnetic radiation from the Sun. This increases the ionisation of the ionosphere, which changes the degree of reflection and absorption of radio waves. For example, the lower layer absorbs more of the lower frequency radio waves and this can create interruptions in radio communications on Earth. A solar flare is shown in figure 3.7.

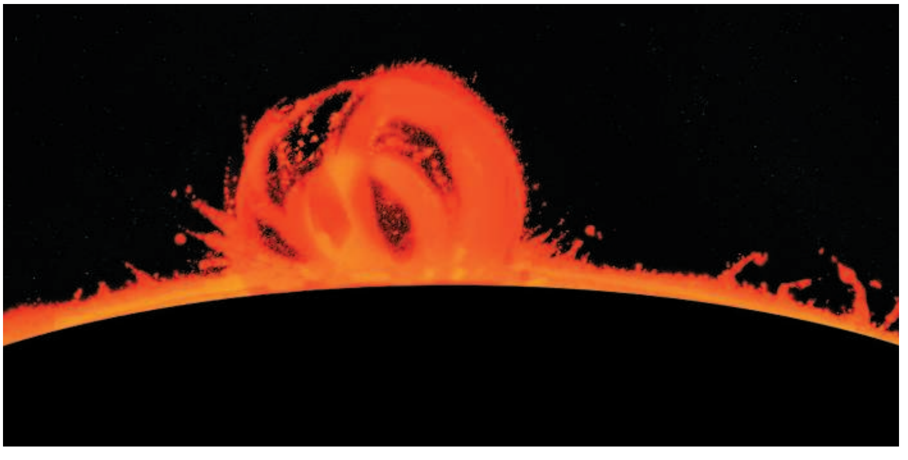


Figure 3.7 *Sudden bursts of energy from solar flares disturb the levels of electromagnetic radiation in the ionosphere and this can create problems with radio transmissions and communication.*

The communication spectrum

The part of the electromagnetic spectrum that can currently be used in communication technologies is limited. This restricted range of frequencies has caused governments to limit the bandwidth over which certain communication devices can operate. There are a number of reasons for keeping the major users of bandwidth at a distance from each other on the spectrum, for example:

- to avoid the problems with interference. Different technologies need different bandwidth separations. For example, television stations do not broadcast on frequencies closer than 5 MHz, and FM radio stations require a 200 KHz separation.
- to provide equity for users. The communications industry is competitive and bandwidth is the subject of intense competition.
- to enable communications to form part of the safety infrastructure
- to ensure communications systems can develop to a standardised plan
- to allow for new technologies to be developed that may require spectrum bandwidth.

The mobile phone is a good example of the way in which new technologies can take up bandwidth. 'Mobiles' use part of the spectrum that was previously left without communication traffic. And 20 years ago, no-one anticipated the volume of mobile phone 'traffic' that we see today.

Governments all over the world restrict the bandwidth available for different users and license that bandwidth. This enables governments to ensure that users have certainty that they can broadcast or communicate over a certain bandwidth and confidently invest in developing technology that is able to use that bandwidth for communication.

Current bandwidths for Australia are given in table 3.1.

Table 3.1 **Bandwidths for communications in Australia**

AM radio	500 to 1500 kHz
Short-wave radio	1.605 to 54 MHz
TV and FM radio	54 to 1600 MHz
Microwaves and radar	1.6 to 30 GHz
Telemetry and experimental	30 to 300 GHz
GSM mobile phones	900 and 1800 MHz
3G mobile network	2100 MHz

Infra-red radiation and light are also used for communication over long distances via optical fibres. In this case, their use is not restricted by the range of spectrum available because they are utilised within enclosed systems where penetration and attenuation are the issues that restrict the spectrum range used.

PHYSICS FACT

The importance of ozone

Ozone (O_3) is a gas that is important in absorbing harmful UV rays. It is a critical component of the Earth's atmosphere because, during its constant production and breakdown, it absorbs harmful solar ultraviolet radiation at wavelengths less than about 320 nm. Ozone is found in relatively higher concentrations (compared to other gases) 10–50 km above the Earth's surface than it is at the Earth's surface. The greatest concentration of ozone is, however, 15–30 km above the surface of the Earth in the layer of the atmosphere called the stratosphere.

Because of the presence of ozone in this stratospheric layer, and ozone's strong absorption of solar UV, it is virtually impossible for UV rays between 200 and 300 nm to penetrate to the Earth's surface. UV with a wavelength of 290 nm is 350 million times weaker at the Earth's surface than at the furthest point of the atmosphere. At 40 km above the surface, about 50 per cent of UV with a wavelength of 290 nm has been absorbed by the ozone.

UV radiation is divided into three types based on its wavelength:

- UVa (320 to 400 nm)
- UVb (280 to 320 nm)
- UVc (200 to 280 nm).

The shorter wavelength radiation has the most energy, but is the first UV absorbed. UVc is absorbed by the small concentrations of ozone high in the atmosphere. The lower layers of the atmosphere then effectively absorb any that does manage to penetrate the upper atmosphere so that practically none reaches the Earth's surface. UVb is mostly absorbed (about 90 per cent or more) by the upper atmosphere layer called the troposphere and the rest is largely absorbed when the UVb passes into the lower atmosphere. About half of the UVa from the Sun is absorbed by ozone or scattered before reaching the troposphere. More is absorbed in the lower atmosphere. Even so, there is enough of this relatively high-energy radiation penetrating through to the surface of the Earth to cause sunburn and skin cancer. Without the ozone in the atmosphere this would be much worse and would affect many more people.

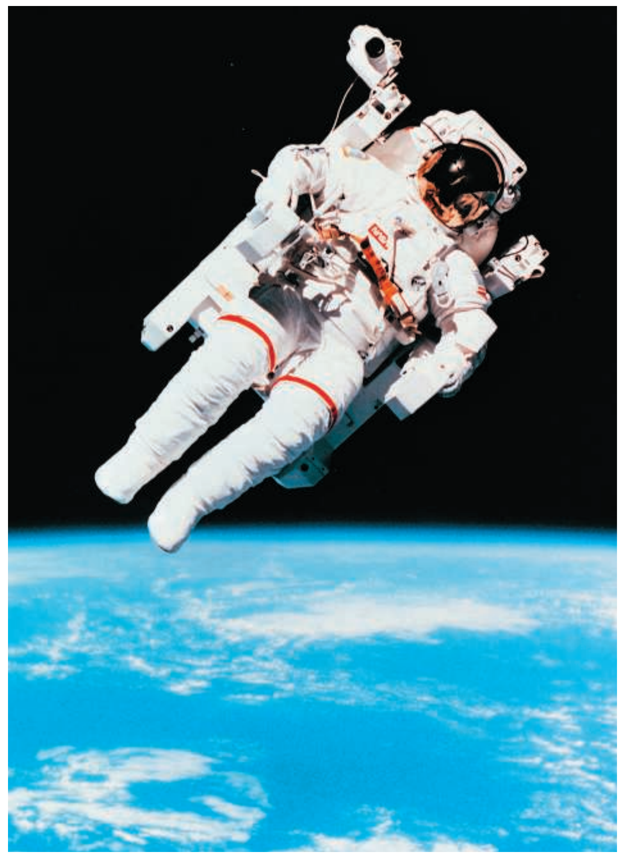


Figure 3.8 ***Astronauts who walk in space or on the Moon require a strong UV filter in the face shield of their spacesuits. This prevents damage to their skin and eyes from the increased UV radiation levels in atmosphere-free and ozone-free environments.***

Light and the atmosphere

The intensity of visible light is affected by the atmosphere. Research telescopes are always placed at relatively high altitudes to avoid the thickest part of the atmosphere closest to sea level.

White light is made up of all the colours of the visible spectrum. The visible spectrum of electromagnetic radiation extends from red light, at around 700 nm wavelength, to violet, at around 400 nm. However, not all of that radiation is equally successful at penetrating the atmosphere.

Evidence for this uneven penetration is seen when the path that the light must travel through the atmosphere is at its longest; that is, at sunrise and sunset. Light must travel further and penetrate more of the atmosphere at sunrise and sunset when the Sun is low — it is travelling 'across' the atmosphere, not down through it as it would around the middle of the day (see figure 3.9). The Sun looks red at sunrise and sunset. Red light is at the low-frequency, longer wavelength end of the visible spectrum. Violet light is at the high-frequency, shorter wavelength end of the visible spectrum. Therefore, visible light of low frequency penetrates the atmosphere more successfully than visible light of high frequency.

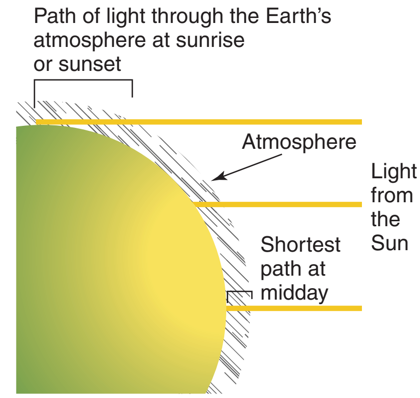


Figure 3.9 *Visible light must travel further through the atmosphere at sunrise and sunset than at midday.*

3.3 ELECTROMAGNETIC RADIATION AND THE INVERSE SQUARE LAW

The greatest advantage of communication technology that uses electromagnetic waves is that the communication can be made extremely rapidly over long distances. The greatest problem with long-distance communication is that the signals decrease in strength the further they are from the source. Examples of this happening include:

- the intensity of the signal indicator on a mobile phone drops off as you move away from the transmitting antenna
- the signal of your favourite radio station can no longer be received as you move a long distance away from the location where you normally listen to the station
- you cannot get the signal from neighbouring cities' television stations as clearly as you can receive the signal from nearby stations (television broadcasts have a maximum broadcast range of 120 to 160 km)
- the further you are from a light source, the lower the intensity the light from that source appears to be. If you are 1000 metres from a car's headlights, the light from them appears dim. The lights illuminate an area distant from them much less brightly or with lower intensity than the area immediately in front of the headlights.

Attenuation is the fall off in energy that occurs as a wave passes through a medium.

3.1

Modelling light intensity versus distance

This decrease in the strength of the signal or light is known as **attenuation**. To reduce attenuation effects in long-distance communication the electromagnetic waves need to be either sent out as a very large strength signal initially or the signals need to be amplified at repeater or booster stations along their path. For very weak signals the receiver may collect a signal over a large area and focus the signal to increase its strength, such as is accomplished with satellite dishes and radio signals from space. Interestingly, the satellites involved in communications networks are themselves booster stations, receiving relatively weak signals which they amplify before forwarding on.

The relationship between distance from the source and strength or intensity of the electromagnetic-wave signal is clear. Increasing the distance between the receiver and the source results in a decreased intensity of the received signal. The relationship between intensity drop off and distance from source is an example of the inverse square law.

The inverse square law

The inverse square law as applied to electromagnetic waves and distance from their source is $I \propto \frac{1}{d^2}$. That is, the intensity of the signal varies inversely with the square of the distance. If the receiver is twice the distance from the source, the intensity is reduced to a factor of $\frac{1}{2^2}$ or $\frac{1}{4}$ of the original intensity. If the distance from the source of the radiation is increased by a factor of 3, the intensity is decreased to $\frac{1}{3^2}$ or $\frac{1}{9}$ of the original intensity. This is illustrated in figure 3.10.

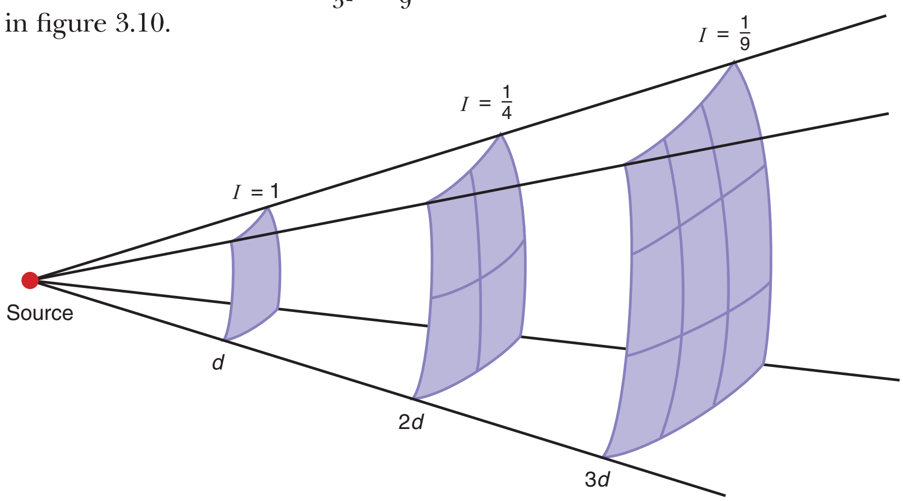


Figure 3.10 *The inverse square law states that the intensity of electromagnetic radiation is inversely proportional to the square of the distance from the source of the radiation. When the distance is doubled, the intensity decreases to one fourth of the original value, and so on.*

3.2

Investigating the inverse square law

The easiest type of electromagnetic radiation to observe is light. Light intensity (illuminance) is measured in units called lux (lx) using a light meter. To visualise the unit 'lux', consider the following:

- On an average day sunlight ranges from 32 000 to 100 000 lx.
- Rooms are lit to about 500 lx by an incandescent light for a comfortable light level.
- Moonlight lights up an area with an intensity of about 1 lx.

Using the inverse square law, if you were one metre from a light source where the light intensity was 16 000 lx, then:

- at two metres the light intensity would be: $\frac{1}{2^2} \times 16\,000$ lux = 4000 lx
- at three metres the light intensity would be: $\frac{1}{3^2} \times 16\,000$ lux = 1778 lx
- at four metres the light intensity would be: $\frac{1}{4^2} \times 16\,000$ lux = 1000 lx.

3.4 MODULATION OF WAVES TO TRANSMIT INFORMATION

All waves carry energy from place to place. A wave that carries exactly the same amount of energy continuously does not carry information. For a wave to be useful and carry a message, the wave must vary. There are two simple ways to vary a wave to add information to it. You either vary the frequency (and hence wavelength and energy) of the wave or you vary the amplitude (and hence energy) of the wave. When you speak, you add information to the sound waves that are your voice by varying their amplitude and the frequency. These variations give the signal you know as language.

Modulation is the process of changing the amplitude or frequency of a wave to add a signal.

Bandwidth describes a series of adjacent frequencies forming a band within the spectrum.

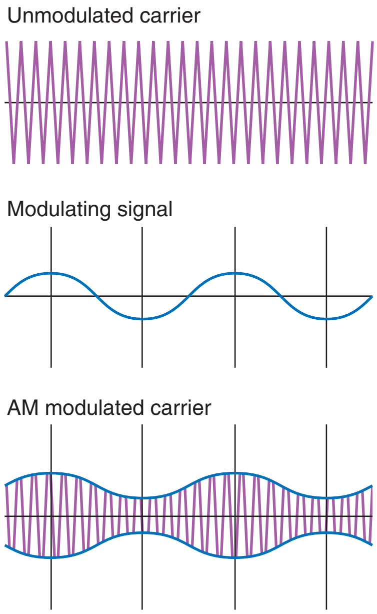


Figure 3.11 Amplitude modulation of a carrier wave

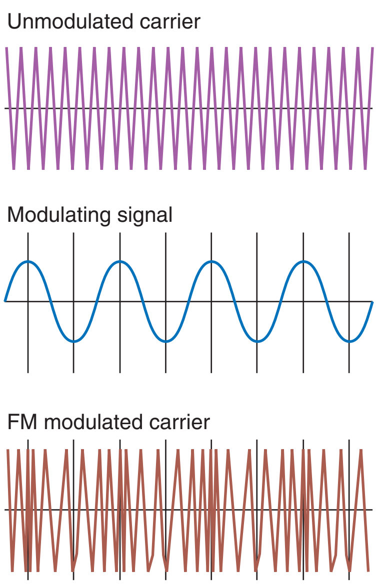


Figure 3.12 Frequency modulation of a carrier wave

The changes in amplitude and frequency of the sounds coming from your larynx have added information to the signal of your voice by modulating the sound wave. Electromagnetic waves can carry information but, like your voice, the information must be added to the waves. The process of adding the signal information to an electromagnetic wave is called **modulation**.

As with your voice, the modulation of an electromagnetic wave involves the alteration of its frequency or amplitude in order to add the information. Radio or microwaves are most commonly used, although the technology of fibre optics is now in favour for sending large quantities of information. The modulated wave (with a varied amplitude or frequency) that makes up the signal must then be converted back into information you can use. That process is called demodulation.

Modulating a radio wave

A radio-wave signal occupies a **bandwidth** of frequencies. This means that the transmitted electromagnetic wave is using a number of frequencies that lie next to each other rather than a single frequency of radio wave. In the middle of that bandwidth is the carrier wave. The carrier wave does not carry any information; it is a by-product of the radio-wave transmitter and acts as the central frequency to which you tune your receiver. The message signal is added to that carrier wave. This is done by superposition of a signal wave onto the carrier wave.

When a radio wave is to be modulated to carry a signal, the amplitude of the wave will be modulated if the information is being sent on an AM (amplitude modulation) band or the frequency will be modulated if the information is being sent on an FM (frequency modulation) band.

Amplitude modulation

Amplitude modulation is commonly used in broadcasting AM radio. The signal is added to an AM carrier-wave radio signal by changing the strength of the signal in a way that corresponds to the information carried by the broadcast. The AM signal remains constant in frequency bandwidth but the amplitude of the wave varies. The variation in the amplitude of the wave is decoded by a radio receiver to produce the signal which is amplified by internal circuitry and converted to the sound signal you hear.

Figure 3.11 illustrates that amplitude modulation is an example of superposition of the modulating signal that carries the message onto the carrier wave.

Frequency modulation

A frequency-modulated (FM) radio transmission means that the signal part of the wave has been added to the carrier wave to vary the frequency of the wave. As shown in figure 3.12, frequency modulation is also accomplished with superposition of a signal wave onto a carrier wave.

A limiting circuit in the radio receiver removes any amplitude variations that occur during transmission of the radio signal and keeps the amplitude of the received wave near constant. The signal is converted back into sound by a discriminator circuit.

Advantages and disadvantages of FM and AM for broadcasting

FM broadcasts are now used more commonly than AM broadcasts for transmitting music. Most of the natural and artificial sources of radio noise, called static, are AM in nature. The effects of 'noise' are reduced in FM radio broadcasts by the limiting circuit in the receiver. This is because the FM radio signal is not dependent on the strength (amplitude) of the signal received, but rather relies on the frequency

Figure 3.13 ***Internet and email capabilities along with the capacity to transmit digital images are now part of mobile phone technology.***

changes to provide the radio signal. It is much harder (although possible) to change the frequency only slightly by interference and superposition. Hence, the music received is closer to that broadcast.

AM radio does have some advantages over FM radio. FM radio signals require a large bandwidth of the limited electromagnetic spectrum; that is, the range of frequencies required to transmit a signal is large. Since the electromagnetic spectrum is limited, the number of FM signals able to be transmitted is limited. AM radio requires a much smaller bandwidth of frequencies for transmission. Because of this, the number of transmissions possible in the AM band is larger.

Modulation of microwaves

The principle of modulating the frequency of electromagnetic waves to carry a signal also applies to microwaves used for transmitting mobile phone signals and mobile internet services. Microwaves are preferred over longer wavelength radio waves for mobile telephone systems because:

- the electromagnetic spectrum is limited and the microwave bandwidth has the capacity available
- microwaves do not spread out as rapidly as radio waves, so more of the transmitted energy makes it to the next receiver dish
- the number of signals able to be transmitted is dependent on the number of frequencies available for frequency modulation. Because the range of frequencies in the microwave transmission range is large, it is possible to send a larger number of signals at once using the same wave. Up to 20 000 telephone calls can be transmitted at once in the microwave band.

The 3G mobile networks send and receive data in the 2100 MHz band. These networks provide access to data, video and the internet on mobile devices such as phones. The 3G networks use a 6 MHz channel carrier width to deliver the higher data transfer rates and increased capacity necessary to deliver these services. Demand for these services will increase in the future and may put pressure on the available spectrum set aside in this band.

Microwaves require a line of sight from one antenna to the next. This means that the proximity of transmitters is controlled by topography. This is less of a problem with the longer radio-wave transmissions. The microwave signals are diffracted by objects larger than their wavelength. As a result, reception in buildings is more difficult for shorter wavelength microwaves. Regular boosts to the signal and more sensitive receivers are required than is the case with radio-wave transmissions.

Microwave transmissions also have their range affected by atmospheric conditions, such as the moisture content of the air. Oxygen also absorbs microwave energy. Moisture and oxygen molecules absorb microwaves in the atmosphere in the same way that water molecules absorb microwave energy in a microwave oven (thus generating heat).

Modulation of visible light

Light modulation is also used to carry signals. The earliest modulation of a light signal by amplitude modulation was probably the helioscope, where an ‘on’ or ‘off’ signal was flashed to communicate over a distance. Laser light of a fixed frequency range will allow transmission of AM signals from a laser transmitter to a receiver. The signal in this case is amplitude modulated because the frequency of light from a particular laser is of a fixed range too small for effective frequency modulation.

Optical fibres are discussed further in chapter 4, pages 61–62.

A device such as the LaserDot® transmitter and receiver can send and receive sound waves transmitted across an open space by an amplitude-modulated laser beam. Such an open-to-air laser device is able to reliably transfer a sound or data signal around 200 m, without the use of fibre-optic cable. Longer distance transmission is less reliable because of the possibility of interference. For that reason, long distance communication by shorter wave electromagnetic radiation, such as infra-red radiation and light, is accomplished along fibre-optic cables to eliminate the chance of interference to the signal.

PHYSICS IN FOCUS

Radar: an application of radio waves

Radar works by sending out evenly spaced pulses of radio-wave energy of a precisely-known wavelength. Electrons moving in an alternating current that has a precisely controlled frequency move up and down in the aerial to create these pulses. If the wave pulses strike an object, such as an aeroplane or car, they reflect back, creating an echo that the radar antenna can detect. Usually the source of the radio waves and the detector are the same aerial. The received radio waves generate an alternating electric current in the aerial. A computer receives the alternating-current information about where the radio-wave echo is coming from, whether the wavelength has been altered and how long each pulse takes to return. It then creates a visible-light track of the moving object on the radar operator's visual display.

Because radio waves travel at the speed of light, the radar must have an accurate, high-speed clock to use the information it gathers. Using such special signal-processing equipment, the radar can measure the speed of the object. It does this by looking at the Doppler shift, or change in wavelength, of the radio waves as a result of being reflected off the moving object, as follows:

- If the object is moving away from the source of the radio waves, the wavelengths will be lengthened as a result of the collision and reflection.
- If the object is moving towards the source, the wavelength of the reflected radio waves is shorter.

The relative change in the wavelength is proportional to the speed of the object either towards or away from the radar antenna.

The Doppler effect also applies to sound waves. When a siren is heard coming from an approaching vehicle, the pitch of a constant frequency appears higher to a stationary observer. When coming from a vehicle travelling away, the pitch appears lower to a stationary observer. The amount of change in the pitch is proportional to the velocity of the vehicle.

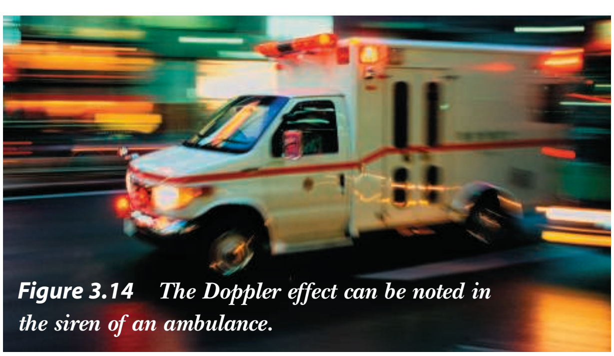


***Figure 3.14** The Doppler effect can be noted in the siren of an ambulance.*

Radar has many applications (see figure 3.15), including:

- the tracking of aeroplane traffic into and out of busy airports
- for police purposes to detect the speed of passing motorists
- mapping of the Earth's surface — and even below the surface of the ocean — by satellites
- tracking satellites and space debris
- military uses, to detect the enemy and to guide weapons
- weather forecasting, to track storms and cyclones
- automatic devices, such as those for opening doors.

***Figure 3.15** Examples of the many uses of radar technology (a) Air traffic controller at airport (b) Police using a radar gun to determine motorist's speed (c) Radar map of Earth's surface*

SUMMARY

- A continuum of radiation types makes up the electromagnetic spectrum. In increasing wavelength order they are: gamma rays, X-rays, ultraviolet rays, visible light, infra-red rays, microwaves and radio waves.
- All electromagnetic waves are produced as a result of alternating electric and magnetic fields, and propagate without the need for a medium in which to travel. All travel at the speed of light (3×10^8 m s^{-1}) in a vacuum.
- The electromagnetic waves most commonly used for communication include: radio waves, microwaves, infra-red rays, visible light and ultraviolet rays.
- The penetrating power of an electromagnetic wave is determined by its interaction with the materials it strikes. In general, longer wavelength radiation is more penetrating than shorter wavelength radiation.
- The drop-off in intensity of electromagnetic radiation is proportional to the inverse of the square of the distance from its source.
- A signal is added to electromagnetic waves in a process called modulation. Modulation can be either through varying the frequency of the wave, for example FM radio, or by varying the amplitude of the wave, for example in AM radio.
- The limitations in the use of electromagnetic radiation are largely due to limitations in the availability of bandwidth of the spectrum suitable for communication. Hence, it is important to utilise the maximum amount of the spectrum and a variety of wave types.

QUESTIONS

1. Describe the effect of the atmosphere on the penetration of very short wavelength electromagnetic radiation such as X-rays from the Sun.
2. Identify the properties that all electromagnetic radiation types have in common.
3. List the electromagnetic wave types in order of increasing wavelength.
4. Identify a communication technology that uses:
   (a) long radio waves
   (b) short radio waves
   (c) microwaves
   (d) infra-red radiation
   (e) visible light
   (f) ultraviolet light.
5. Explain what makes a wave electromagnetic in character.
6. Explain what is meant by the term ‘continuum of electromagnetic radiation’.
7. Astronomers making ground-based observations use radio waves and light waves predominantly to make their observations. Useful information is obtained from all electromagnetic wave types. Explain why the ground-based astronomers only use a small part of the electromagnetic spectrum for observations.
8. Explain why sunburn is such a problem at high altitude even though the temperatures of the environment are much reduced compared to those at sea level.
9. Explain why a telescope observing in the UV range operates best at high altitude and at night.
10. Infra-red radiation is favoured for use in many fibre-optic applications, even though it can carry less information than visible light when used in the same application. Suggest an advantage infra-red radiation may have over visible light for long-distance communication along fibre-optic cables.
11. A student’s results from an experiment to determine the relationship between light intensity and distance from source are recorded in the table below.

    Graph the light intensity against $\frac{1}{d^2}$. Draw a line graph and join your points with a line or curve of best fit. Does the data confirm or approximate the relationship $I \propto \frac{1}{d^2}$? Explain your answer.

LIGHT INTENSITY (lx)	DISTANCE FROM LIGHT (m)	(DISTANCE FROM LIGHT)^2	$\frac{1}{\text{DISTANCE}^2}$
10 000	1	$1^2 = 1$	$\frac{1}{1}$
2 500	2	$2^2 = 4$	$\frac{1}{4}$
1 070	3	$3^2 = 9$	$\frac{1}{9}$
585	4	$4^2 = 16$	$\frac{1}{16}$
390	5	$5^2 = 25$	$\frac{1}{25}$

12. Radar is an application of electromagnetic radiation used to obtain information. Identify the type of wave involved in the use of radar.

13. Describe the operation of radar to observe a moving object, such as an aeroplane, in terms of:
    (a) the energy transfers involved
    (b) the waves involved in the transfers.
14. Explain why it is that, even if a device capable of modulating X-rays and detecting those X-rays was invented, X-rays would still be unlikely to be successful as communication devices.
15. There is a limit to the number of radio stations possible in the AM band. Explain why it is not possible to increase the number of stations beyond the limit.
16. Explain why the ionosphere can be treated as a plane mirror when transmitting short-wave radio, even though it is a curved surface.
17. Some people interested in the search for extra-terrestrial communication argue that we should be looking for pulses of visible light rather than radio waves. Identify the features of visible light that would make it suitable as a communication medium across the vast dimensions of space.
18. Radar uses high frequency radio waves. Explain the advantages of using high frequency radio waves over low frequency radio waves for this purpose.
19. Terahertz waves exist between the infra-red and microwave portions of the electromagnetic spectrum and have the potential to be useful for diagnosis of internal injury in much the same way as X-rays. A barrier to their use is that the technology to effectively produce and detect them has not yet been developed. What might be some of the potential safety advantages of using these rays over X-rays?
20. If terahertz waves could be produced, modulated and detected, would they be able to carry more or less information than microwaves? Justify your answer.

3.1 MODELLING LIGHT INTENSITY VERSUS DISTANCE

Aim

To investigate how the light intensity varies with distance from a point source of light.

Apparatus

round balloon
marker pen
ruler

Theory

The surface of the balloon is an analogy for the energy emitted by a pulsed light source. As the balloon inflates it is mimicking the expanding sphere of light that a single pulse would produce in three dimensions if it could be slowed down and observed from a distance. The expanding surface of the balloon represents the energy distribution of the fixed energy of the light pulse as it changes with time and moves away from its point source. The relationship between the surface area and the energy distribution is an inverse one.

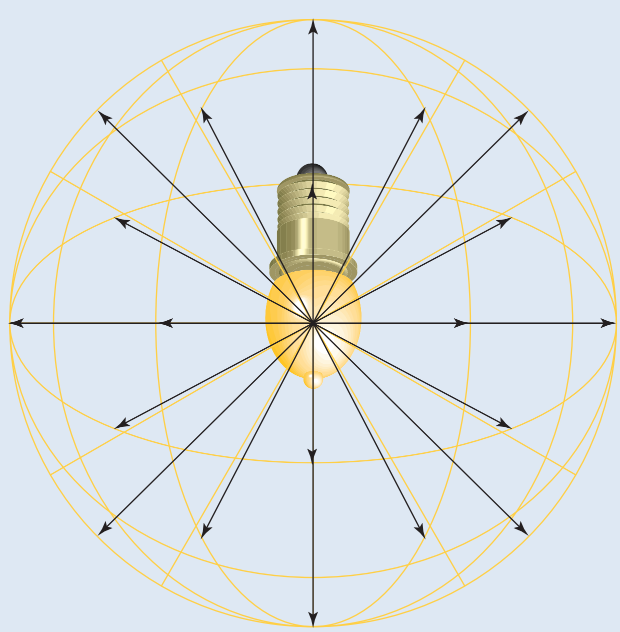


Figure 3.16 *A light globe acting as a point source of light pulses surrounded by a sphere of light ever expanding with time as it travels away from the source*

Method

1. Imagine that a light globe is always in the centre of your balloon. The inflating balloon surface is a representation of a pulsed-light wavefront travelling in three dimensions from the light globe to form a spherical shell of light.
2. Inflate the balloon until it has a diameter of about 10 cm. *Do not tie off the balloon.* Record this as reading 1 in the table below.
3. Holding the inflated balloon, use a marker pen to draw a 1 cm by 1 cm square on the balloon. This is best done where the balloon skin is thickest, usually opposite the inflation tube.
4. Record the area of the square as 1 cm^2 in the space beside reading 1 in the table below. This square represents the energy distribution and hence, intensity of the light from the imaginary globe at the centre of the balloon at that radius from the light source.
5. Inflate the balloon until it has a diameter of about 20 cm. The distance to the centre of the balloon has doubled at this point. This represents a doubling of the distance from the imaginary light source. Record this as reading 2 in column 1 of the table below.
6. Measure the changed size of the square you drew on the balloon. Record the area of the square in the table below, next to reading 2.
7. Inflate the balloon until it has a diameter of about 30 cm. The distance to the centre of the balloon has now tripled. Record this as reading 3.
8. Measure the changed size of the square on the balloon now. Record the area of the square next to reading 3 in the table below.
9. Calculate the value of $\frac{1}{\text{area}}$ and $\frac{1}{(\text{distance units})^2}$. Record those values in the table below in columns 3 and 4.
10. Plot a graph of $\frac{1}{\text{area}}$ representing intensity versus $\frac{1}{(\text{distance units})^2}$.

Results

DISTANCE UNITS FROM THE BALLOON CENTRE	AREA OF THE SQUARE	$\frac{1}{\text{AREA}}$	$\frac{1}{\text{DISTANCE UNITS}^2}$
1	1 cm^2		
2			
3			

Questions

1. Does the relationship shown by your collected data approximate the relationship $I \propto \frac{1}{d^2}$?
2. What happens to the fixed quantity of energy from a point source of light as it spreads out in terms of the amount of energy per unit area?

3.2 INVESTIGATING THE INVERSE SQUARE LAW

Aim

To use a data logger or light meter to investigate the relationship between light intensity and distance from the light source

Apparatus

light globe
a room able to be blacked out
light meter or light metering probe for a data logger
tape measure

Theory

The relationship $I \propto \frac{1}{d^2}$ is an example of the inverse square law. This relationship will be encountered a number of times throughout your physics course. You may use a data logger or light meter to confirm the relationship, $I \propto \frac{1}{d^2}$.

Method

1. Set up the light globe in the centre of a dark room.
2. Point your light intensity probe directly at the light globe, close to the globe (around 0.01 m from it). Record the light intensity or set your data logger to initiate recording readings.
3. Record readings of light intensity at 1 m, 2 m, 3 m and 4 m distances from the light source.
4. Plot your values of light intensity against the $\frac{1}{\text{distance}^2}$. This will give you a graph of light intensity versus the inverse of the square of the distance from the light globe.

Question

Does your analysis confirm this relationship ($I \propto \frac{1}{d^2}$)?

CHAPTER 4 REFLECTION AND REFRACTION OF ELECTROMAGNETIC WAVES

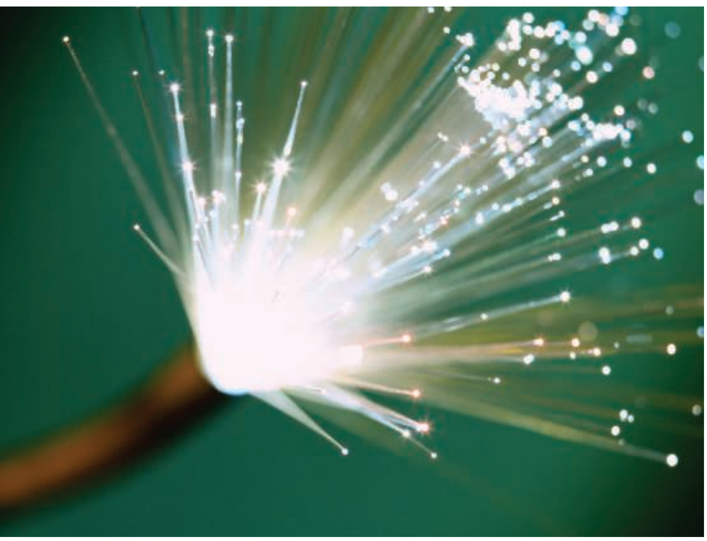


Figure 4.1 *Optical fibres. Transmission of information through optical fibres relies on the phenomenon of total internal reflection of the light or radiation carrying the information.*

Remember

Before beginning this chapter, you should be able to:

- understand the difference between the absorption, reflection, refraction and scattering of light and suggest situations where they occur
- describe some of the common uses of electromagnetic radiation, in particular its use in communication applications.

Key content

At the end of this chapter you should be able to:

- discuss the law of reflection and use it to explain the reflection of waves from a plane surface
- construct ray diagrams indicating the direction of travel of light rays and relate this to a propagating wavefront
- describe how the reflection of light and radio waves is used to transfer information
- present information to show the paths of light rays reflected from plane surfaces, concave surfaces and convex surfaces, and describe an application of reflection for each type of surface
- draw a diagram to show the path of a short-wave radio signal that has been bounced off the ionosphere
- explain that refraction is related to the velocities of waves in different media and outline how this may result in the bending of a wavefront
- perform an experiment to measure the angles of incidence and refraction of light that encounters an interface between the media and show the relationship between these angles
- calculate the refractive index of glass or perspex
- define, apply and solve problems using Snell's Law: $\frac{v_1}{v_2} = \frac{\sin i}{\sin r}$
- identify the conditions necessary for total internal reflection, referring to the critical angle
- outline uses of refraction and total internal reflection in optical fibres
- identify the types of communication data that are stored or transmitted in digital form
- identify the developments of communication technology with reference to current research, and discuss the physical principles involved in CD technology, DVD technology and global positioning systems.

Many of the technologies used in communication rely directly on the properties of electromagnetic radiation called refraction and reflection. These properties have direct analogies in other types of waves, such as water waves and sound.

4.1 THE LAW OF REFLECTION

When a wave is reflected from a flat-surface plane it will obey the *Law of Reflection*. The Law of Reflection states that the angle of the incoming, or incident, wave in relation to a line perpendicular to the reflecting surface (**normal**) is equal to the angle the reflected wave will make with the normal.

The normal is the line that is perpendicular to the reflecting surface at the point where the ray hits it.

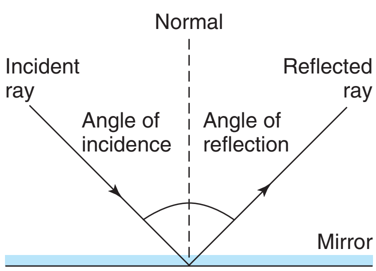


Figure 4.2 ***The ray approaching the mirror (incident ray) is reflected from the mirror surface. The angle of incidence equals the angle of reflection.***

The Law of Reflection can be summarised as: the angle of incidence equals the angle of reflection. Usually the law is represented in diagrammatic form as shown in figure 4.2.

To study reflection, a ray model is useful. The ray model includes the assumption that electromagnetic waves travel in straight lines. Hence, the electromagnetic waves are shown in diagrams as straight lines with arrow heads representing the travel direction. A ray can be likened to an infinitely thin beam of electromagnetic radiation.

The incident ray, the reflected ray and the normal at the point of incidence in figures 4.2 and 4.3 all lie in one plane. That means they can all be drawn as though they lie on one flat sheet of paper. This is shown in figure 4.3.

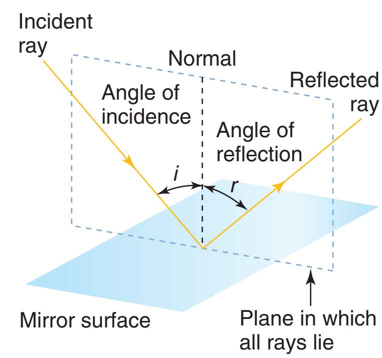


Figure 4.3 ***The incident ray, the normal to the reflecting surface and the reflected ray, all lie in the same plane. For any chosen incident ray, the plane is at 90° to the plane of the reflecting surface.***

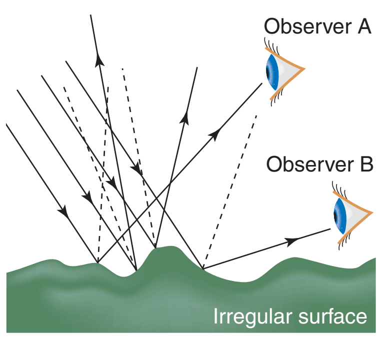


Figure 4.4 ***Reflection from an irregular surface. The dotted lines represent the normal to the reflecting surface at the point of reflection for each incident ray. For each individual incident ray, the Law of Reflection is obeyed. Note that at the points where the incident rays strike the surface, the normals to the surface are not parallel.***

Using the Law of Reflection

Electromagnetic waves can be reflected from smooth, flat, mirror-like surfaces or from irregular or bumpy surfaces. When the waves are reflected from a flat surface, the reflection is regular. When the reflection is from a bumpy or irregular surface, it is diffuse. In both cases, the reflection obeys the Law of Reflection.

If a surface is some distance away from a light source, the light rays that reach the surface are considered to be parallel. When the surface is irregular, each of the parallel incident rays that hits the surface does so at a different angle of incidence. This means that each ray is reflected with a different angle of reflection and the rays are no longer parallel. To two observers, A and B, the reflected rays will not be as intense as the incident rays because the reflection is diffuse (see figure 4.4).

When the surface is regular, the incoming parallel incident rays are all reflected back parallel to each other. For observers, the intensity of the reflected light is greater than that of the same light reflected from an irregular surface (see figure 4.5).

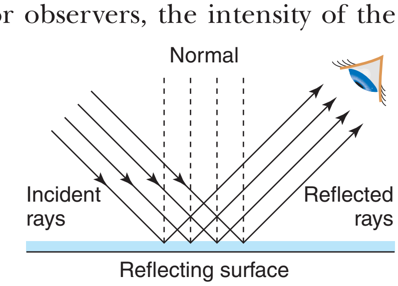


Figure 4.5 ***The incident parallel rays are reflected from the plane surface and remain parallel. An observer sees that the intensity of light is maximised.***

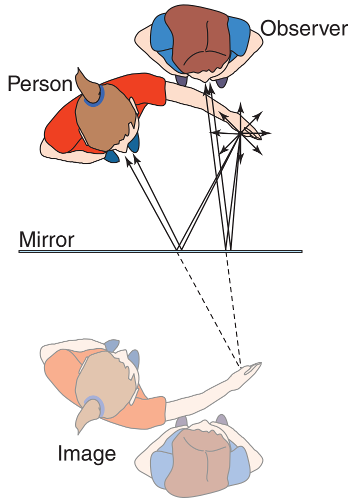
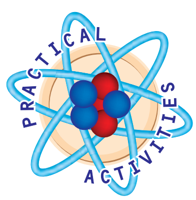


Figure 4.6 ***A person and an observer looking at a reflection in a mirror***



4.1

Using rays to locate an image from a plane mirror

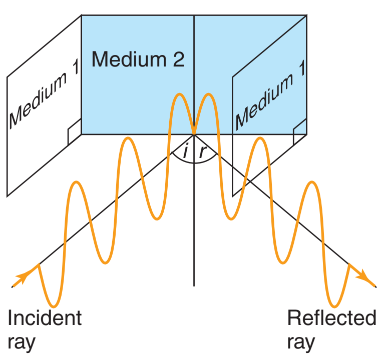


Figure 4.7 ***A phase change occurs on reflection at the interface where medium 1 has a lower refractive index than medium 2.***

The **focus** is the point where all rays from a converging lens or mirror are concentrated. It is also the point where the rays appear to originate after passing through a diverging lens or being reflected by a diverging mirror.

Converging means coming together at a point (the focus).

Light reflecting off water in a pond can produce a mirror-like reflection if the pond is still, or a diffuse reflection if the surface is disturbed. The shimmering effect of light reflected from a disturbed pond-water surface is caused by the constantly changing diffuse reflection of light.

Everyone is familiar with the idea of looking at a plane surface, such as a mirror, and observing a reflection of themselves. The use of the plane mirror for that purpose is probably the most common use of reflection in our daily lives. The Law of Reflection allows you to explain the properties of the image formed by reflection. Those properties include:

- the size of the image
- the nature of the image
- the apparent location of the image
- the orientation of the image compared to the original object.

Images can be classified as being real or virtual in nature. A real image is one that can be projected onto a screen. A virtual image can be seen but cannot be projected onto a screen.

Consider the situation shown in figure 4.6 where a person and an observer are looking into a mirror. When the person stands in front of a plane mirror and looks at their reflection, every point on their body reflects light. The person's body acts as a source of light waves that travel towards the mirror. Some of these waves will strike the mirror and reflect back towards the person's eye, where the light waves are focused to form an image. The image appears as though it is coming from an identical person behind the mirror. Consider the person's hand and its reflection. Every point on the person's hand has a corresponding image point that appears to be behind the mirror. These points assemble to recreate the appearance of that person's hand. An observer looking into the mirror (as in figure 4.6) may see rays that reflect off the mirror that do not enter the person's eye. The image they see will therefore be different, but again will appear to come from behind the mirror. The apparent position of an image of an object in a plane mirror can be located using rays and the Law of Reflection in a geometrical construction. The procedure is outlined in practical activity 4.1 (page 70).

Reflection and wave phase change

When a wave is reflected from the surface of a material having a higher index of refraction, the wave undergoes a phase change of 180°. The reflected wave is out of phase by a half cycle compared with the incident wave. This is the case when light travelling through air is reflected from a glass mirror.

When light is reflected from the surface of a material with a lower index of refraction, there is no phase shift. Note that the frequency of the reflected wave remains unchanged compared with the incident wave. You will learn about refractive index later in this chapter on pages 57–58.

Reflection from curved mirrors

You probably recall seeing curved mirrors in many applications in daily life. Two of the most common applications encountered are the car headlight, shown in figure 4.8, and the reflector for a light globe in downlights or torches. The curved mirrors concentrate the initially parallel light rays to a point known as the **focus**, and are therefore called **converging** mirrors. If the light rays originate from a point source at the focus, such as in a car headlight, they will be parallel after being reflected. These curved mirrors are silvered on the concave side and are known as concave mirrors.

eBookplus

eModelling: **Model of a concave mirror** doc-0055

Rays from a **diverging** mirror appear to come from a focal point behind the mirror and are spread out by the mirror.

Other curved mirrors are silvered on the convex side. They are commonly used to give a wider field of view in car rear-view mirrors and in security and safety situations such as at sharp bends or in shopping centre stores where shoplifting is an issue. These mirrors cause the parallel rays incident on their surface to be reflected as though they diverge from a focus behind the mirror. Hence, they are known as convex or **diverging** mirrors. Safety mirrors are shown in figure 4.9.

Figure 4.8 *A concave or converging mirror used as a reflector in a car headlamp*

eBookplus

Weblink: **Concave mirror applet**

Figure 4.9 *Convex mirrors used as safety mirrors at a worksite*

There are many types of curved mirrors that are designed for specific uses. The most common type of curved mirror is the spherical mirror. Spherical mirrors are easy to make. Spheres of glass are blown by glass blowers, cut into hemispheres, then silvered on one side to make a concave or convex spherical mirror. The terminology to describe these types of mirror is used because they are originally made from a sphere.

To accurately describe how rays are reflected and how curved mirrors form images, we use certain terms. These include:

- the sphere centre, C, shown in figure 4.11
- the pole of the mirror, P, as shown in figures 4.10, 4.12 and 4.13 (a ray projected through the centre of curvature would be bounced back along the same path from a pole point)
- the principal axis, which is the radius of the sphere (from which the mirror may have been made) that passes through C and P.

When ray diagrams are drawn, it is conventional to draw the curved surface as a line. The reflection of the ray is then drawn as though it has come from this straight line rather than a curve.

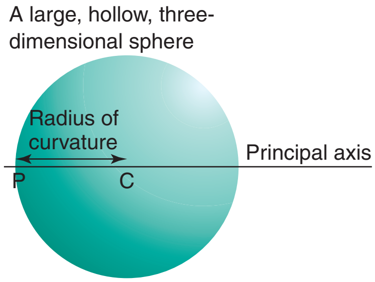


Figure 4.10 *A spherical mirror forms part of a sphere. The distance PC is the radius of curvature of the mirror. PC is the radius of the sphere from which the mirror was made. P is the pole, or central point, of the mirror that would lie on the diameter of the sphere. The diameter of the mirror is called the principal axis of the mirror.*

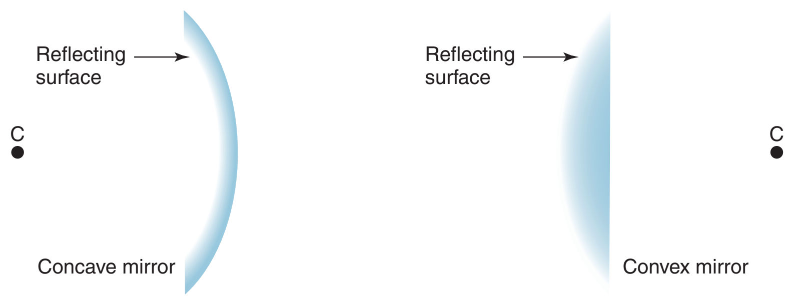


Figure 4.11 *Concave and convex mirrors showing the reflecting surfaces*

When light rays strike a concave, spherical-mirror surface in a direction that is parallel with the principal axis, they are reflected so that they converge to a single point called the focus. This point is the focal point of the mirror. The focus is labelled F in figure 4.12. The focus is at a distance equal to half the centre-of-curvature radius of the mirror, or a distance of half the radius of the sphere from which the mirror could have been made.

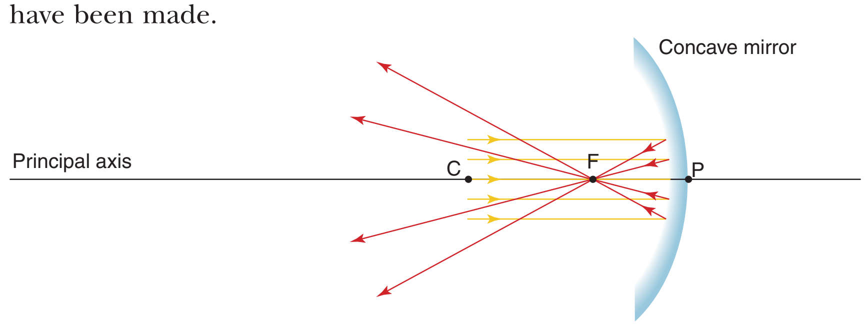


Figure 4.12 *A concave mirror, also called a converging mirror because the reflected rays converge at the focus in front of the mirror*

When light is reflected by a convex mirror, any incident rays hitting the mirror (with a direction parallel to the principal axis) are reflected and diverge as though they originate at a point behind the mirror. In a similar way to finding the apparent location of the image produced by a plane mirror, the point where these reflected rays appear to diverge behind the convex mirror is the focus.

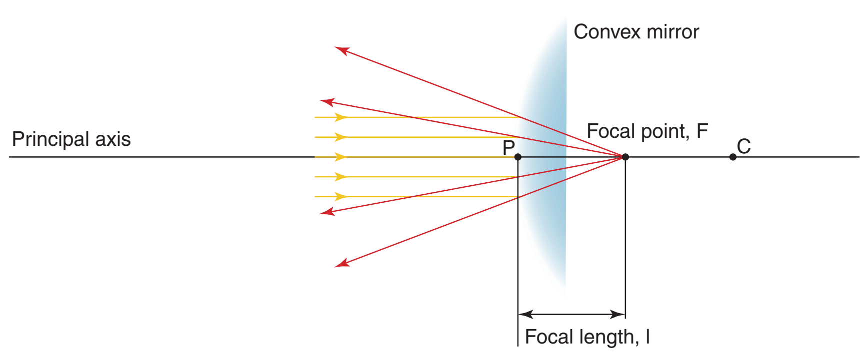


Figure 4.13 *A convex mirror showing the parallel incoming rays diverging — apparently from a focus, F, located behind the reflecting surface of the mirror. The distance between the focus and the pole of the mirror, P, is the focal length. Note that, because the focus is behind the reflecting surface, the convex mirror can never produce a real image, although an observer looking into the mirror can see the virtual image it produces.*

Note that, for both the concave and convex spherical mirrors, the focus of the mirror, F, is halfway between P and C.

PHYSICS FACT

Finding the image produced by a curved mirror

You can use ray tracing to determine the position and nature of an image formed by a curved mirror. We will start with the concave mirror, but you will see that the principles are the same for the convex mirror.

One important principle that helps in locating images, or the position of the object that created the image, is that the paths followed by light rays are reversible. This is the Principle of Reversibility of Rays. This principle is used extensively in reflectors for car headlights. The filament of the light globe is placed at the focus of the curved reflector and the result is an emergent reflected beam of light that is parallel. This feature enables the light from the headlights to be concentrated forward.

The nature of an image formed by a mirror is determined by constructing any two of the following four rays using a scale diagram. It is vital that accurate drawings are made to ensure the position and nature of the image is determined.

PRACTICAL ACTIVITIES

4.2

Reflection in concave and convex mirrors

1. The first ray to be drawn travels in a direction that is parallel to the principal axis and is from the top of the object to the mirror surface. After reflection, this ray passes through the focus, or should be drawn as if it comes from the focus if the mirror is convex.
2. The second ray should be drawn from the top of the object through the focus, and on reflection becomes parallel to the principal axis.
3. The third ray is drawn towards the pole of the mirror and is reflected such that the angle of reflection it makes with the principal axis is identical to the angle of incidence the ray makes with the principal axis. However, the angle of incidence is on the opposite side of the principal axis. This ray obeys the Law of Reflection at the pole of the mirror where the principal axis is the normal ray to the mirror at this point.
4. The fourth ray travels from the top of the object through the centre of curvature to the mirror. This will fix the top of the image at the mirror. Such a ray reflects back on itself because it is incident normally on the mirror's reflecting surface.

The bottom of the object lies on the principal axis. A ray projected along the principal axis is always projected back on itself. That ray's projection behind the mirror defines the image bottom.

The intersection of these four rays defines the position and the nature of the image formed by an object on reflection from the mirror. The nature of the image can be predicted based on the position of the object and the mirror type.

For concave mirrors, if the object is:

- inside the focal length of the mirror, the image will be virtual, upright and enlarged (as shown in figure 4.14)
- between the focal length and the centre of curvature, the image will be real, inverted and enlarged
- beyond the centre of curvature, the image will be real, inverted and diminished (as shown in figure 4.15 on page 52).

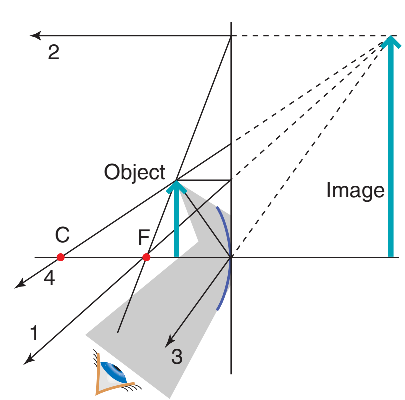


Figure 4.14 *Locating an image where the object is closer to the mirror than the focal length of the mirror. The image formed in this case is virtual, enlarged and upright.*

(continued)

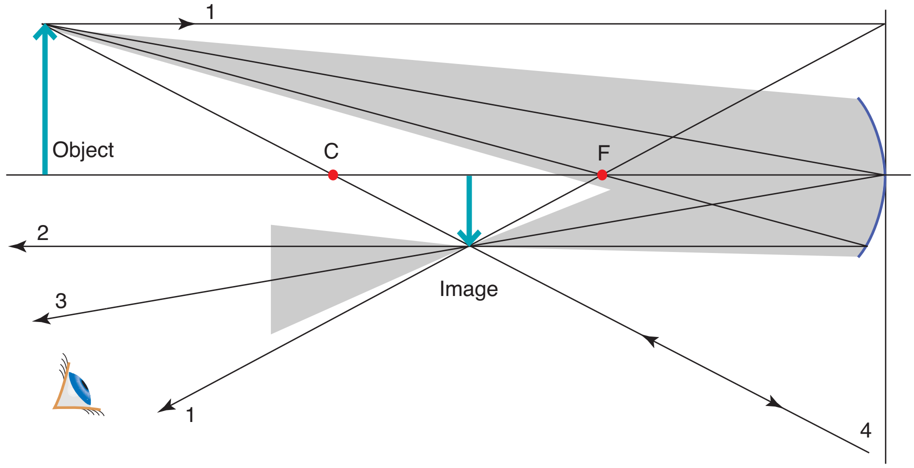


Figure 4.15 ***Locating an image of an object reflected from a concave mirror where the original object is beyond the focal length of the mirror. The image formed is inverted (upside down), and in this case is diminished (smaller than the object).***

If the object is at the focal length, it is a special case where the image actually forms at infinity. If the object is at the centre of curvature, the image that forms is real, identical in size to the original object and inverted.

For convex mirrors, if the object is anywhere in front of the mirror, it produces an image that is upright and diminished in size. The further the object is from the reflecting surface the smaller it will appear. The image will always be virtual and will appear to be behind the reflecting surface.

Determining the image position and character for a reflection of an object by a convex mirror can be accomplished by using any two of the following three rays.

1. The first ray is drawn as though it passes through the centre of curvature of the mirror and reflects back on itself. If the reflected ray is traced back it will define the top of the upright image behind the mirror.
2. The second ray is projected parallel to the principal axis, strikes the mirror and is reflected as though it originated at the focus behind the mirror. If the reflected ray is projected back, it passes through the top of the image behind the mirror.
3. The third ray is drawn as though it will pass through the focus and is reflected back from the surface parallel to the principal axis.

Any two rays intersecting will define the top of the image. The bottom of the object is the principal axis. A ray projected along the principal axis always gets projected back on itself. That ray's projection behind the mirror defines the image bottom. Figure 4.16 represents this method of determining the nature of an image.

The projected rays are broken lines to indicate that the light doesn't actually travel from these points behind the mirror. The image is:

- virtual
- upright
- diminished.

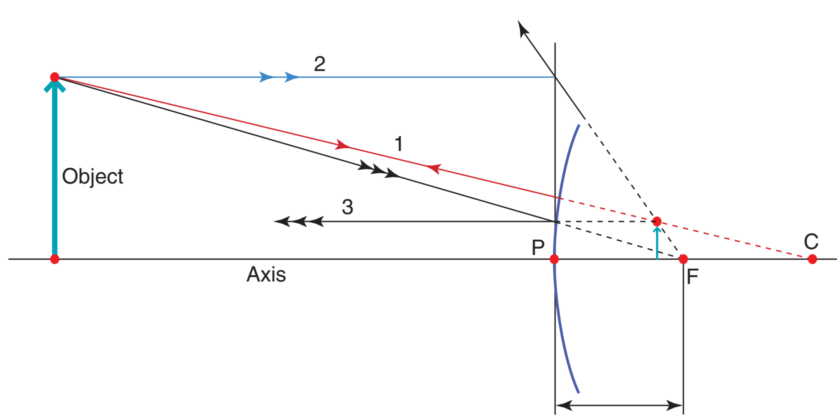


Figure 4.16 ***Determining the image position and character of a reflection of an object by a convex mirror***

4.2 APPLICATIONS OF REFLECTION

The astronomical telescope

The telescope of the type shown in figure 4.17 was invented by Isaac Newton (1642–1727) in the late seventeenth century. It is called the Newtonian reflecting telescope.

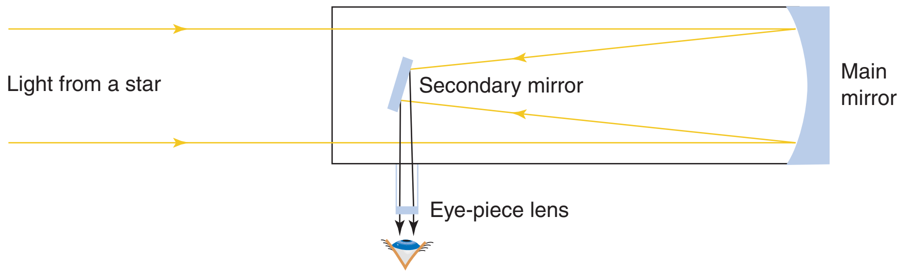


Figure 4.17 *A Newtonian reflecting telescope. The main light-collecting mirror is either a concave spherical mirror or a parabolic mirror. The parabolic mirror has the advantage that it tends to focus the light collected by the mirror to a point more accurately than a spherical mirror. The secondary mirror is a plane mirror.*

Most large astronomical telescopes are Newtonian telescopes, or a variation of them such as the Cassegrain telescope. The Cassegrain telescope was invented in 1672 by a Frenchman, N. Cassegrain. It differed from the Newtonian telescope by having a hole cut in the concave main mirror. The reflected light from the main mirror was reflected back to the eye piece by a small convex or hyperbolic mirror, as shown in figure 4.18. Because the Cassegrain telescope has a different arrangement of reflecting surfaces, it has a shorter barrel length than equivalent focal-length Newtonian telescopes. The image is also viewed in the direction of the object rather than perpendicularly to it as with the Newtonian telescope.

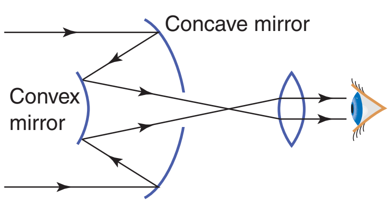


Figure 4.18 *The mirror assemblage and ray path in a Cassegrain telescope*

The main mirrors in both the Newtonian and Cassegrain telescopes are typically coated with aluminium and are front-silvered, concave, parabolic mirrors. Parabolic mirrors have the advantage over spherical mirrors of being able to eliminate spherical aberration (see figure 4.20 on page 54). Spherical aberration is the distortion seen when we view things with a single convex-lens system such as a magnifying glass. It is where the image seen at the centre is clear and focused, but near the edge of the glass the same object appears distorted and out of focus. This effect produces an unclear image.

Two of the largest reflecting optical telescopes located on Earth are the twin Keck telescopes. Their main mirrors are 10 metres in diameter and they are located at high altitudes (4000 m). The Anglo-Australian Telescope at Siding Springs, near Coonabarabran, New South Wales, has a 3.9-metre diameter main mirror.

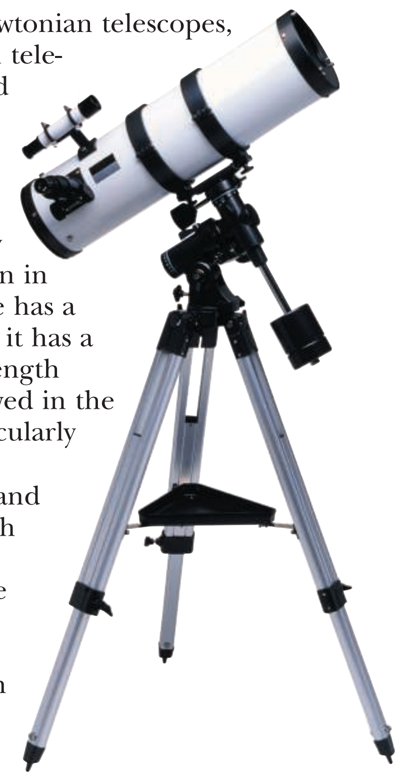


Figure 4.19 *A commercially produced Newtonian telescope that is often popular with amateur astronomers. The small refracting telescope fitted to the top of the reflecting telescope is used to find the astronomical object being observed. These telescopes produce inverted images of the objects observed but this is unimportant to the astronomer.*

The advantage of a large diameter is that the greater light-gathering power enables the observation of fainter astronomical objects; and the larger the diameter of the mirror, the better the resolution of the telescope.

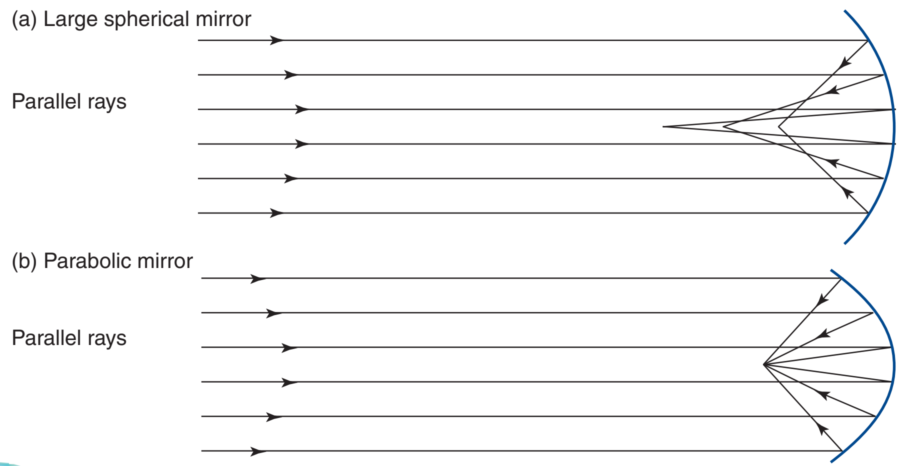


Figure 4.20 ***Large, spherical mirrors focus the parallel rays that hit near the pole of the mirror at a slightly different place than the parallel rays that reflect from near the edges of the mirror. This is known as spherical aberration. Parabolic mirrors do not show spherical aberration and focus all rays at the one focus.***

Figure 4.21 ***A torch with an adjustable parabolic reflector***

Torches and driving lights

Torches and driving lights of superior quality are often advertised as having parabolic, concave-mirror reflectors. In some torches, the ability to move the filament from the focus means the torches are able to produce a spot or flood beam. Figure 4.21 shows a torch with a parabolic reflector. When the beam is adjusted to 'spot' the filament is at the focus. When the beam is set to 'flood' the filament is beyond or before the focus of the reflector mirror.

The satellite dish

Satellite dishes receive weak radio signals from satellites in space or detect naturally emitted radio frequencies from astronomical objects. These signals are received by the dish as parallel rays that hit the dish surface. A large, concave, parabolic reflecting dish collects the weak-intensity signals, reflects them and focuses the signals to a receiver aerial at the focus of the satellite dish. This increases the strength of the signals received. Figure 4.22 shows a satellite dish and its receiver aerial located at the focus.

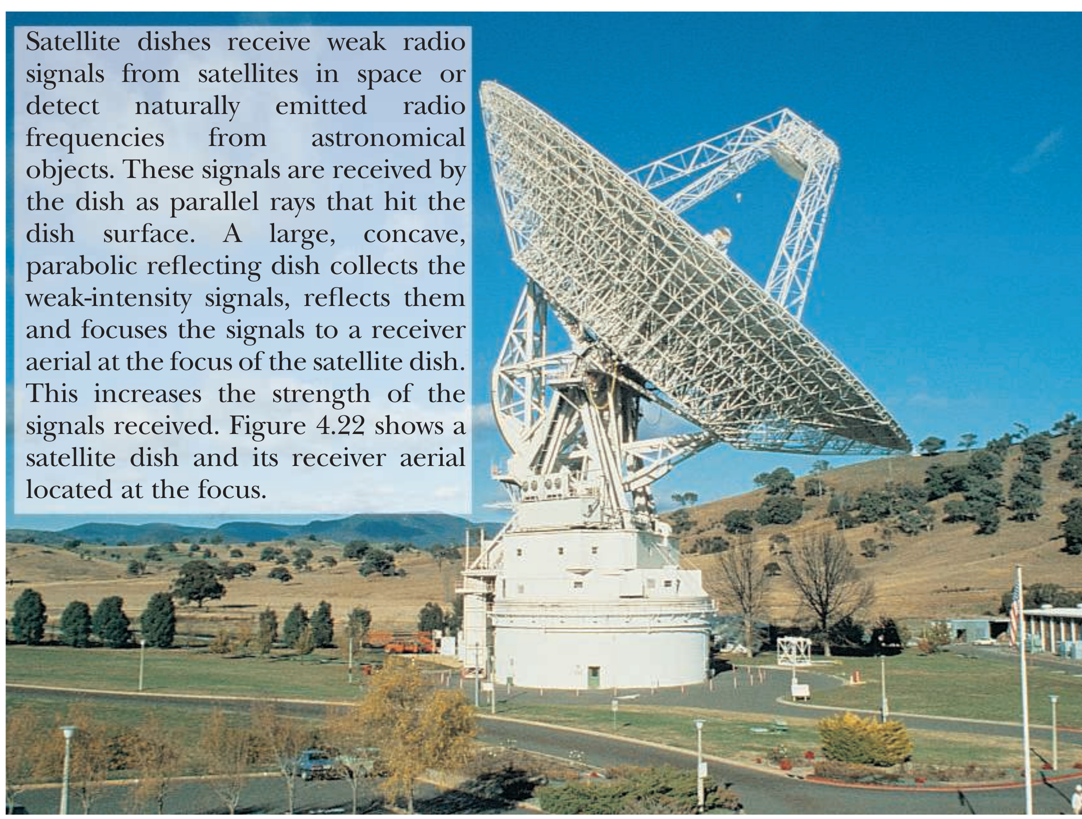


Figure 4.22 ***Radio telescopes collect incoming, parallel, weak radio waves from distant astronomical objects or satellites and concentrate the signal by reflecting it to a single focus where an aerial is located.***

Using reflection in communication: radio waves

As we have seen, radio waves are used in many modern communication technologies. Some of these radio waves are known as sky waves, others are called space waves and yet others are called surface waves.

We saw in chapter 3 that part of the Earth's atmosphere at high altitude is ionised (the atoms and molecules have become charged by losing or gaining electrons) — producing a layer we refer to as the ionosphere. When generated on the surface of the Earth, surface radio waves do not usually reach the ionosphere. Space waves penetrate right through the ionosphere and are used to communicate with satellites and the International Space Station in a direct-line-of-sight manner. Sky waves are special because they bounce off the ionosphere.

Waves in the sky-wave radio band have high frequencies and short wavelengths. When waves of that frequency come from outer space they are unable to penetrate to the surface of the Earth. If the waves are generated on Earth, they are bounced by the ionosphere towards the Earth's surface where they are bounced again back towards the ionosphere. Because of the high altitude of the ionosphere, its curvature is not that large. As a result, the waves are essentially bouncing off a plane surface and so obey the Law of Reflection. In this way, the short-wavelength radio waves can literally bounce their way around the globe from a single transmitter, as shown in figure 4.23. This mode of transmission can be likened to a rubber superball that is thrown forward towards the ground where it bounces off a floor and a low roof multiple times on its forward passage.

Radio Australia uses a wide range of communication technologies to deliver its programs to audiences in an area stretching from Hawaii in the east to Bangladesh in the west, across the Pacific and South-East Asia and up to Japan and Korea. Formats include short-wave radio, satellites, audio CDs, multimedia internet sites and CD-ROMs. Short-wave transmission from sites in Australia, Singapore, Taiwan and the northern Pacific ensure optimal reception of its programs in Asia and the Pacific. Selected programs are also available via satellite or local broadcasters in Europe, North America and Africa.

The ionosphere has one more quirk. You may recall from chapter 3 that the ionosphere actually rises at night when the incoming solar radiation level decreases. This higher ionosphere means that the range of short-wavelength radio waves increases at night as they bounce off the higher ionosphere. This is the reason why short-wave radio broadcasts from distant transmitters unable to be heard during the day can be heard at night.

Most national radio programs such as Radio Australia are designed to be heard at great distances overseas and are purposely broadcast in the short-wave band to enhance their coverage (see the note at left). Short-wave transmissions are often received with much greater clarity at night.

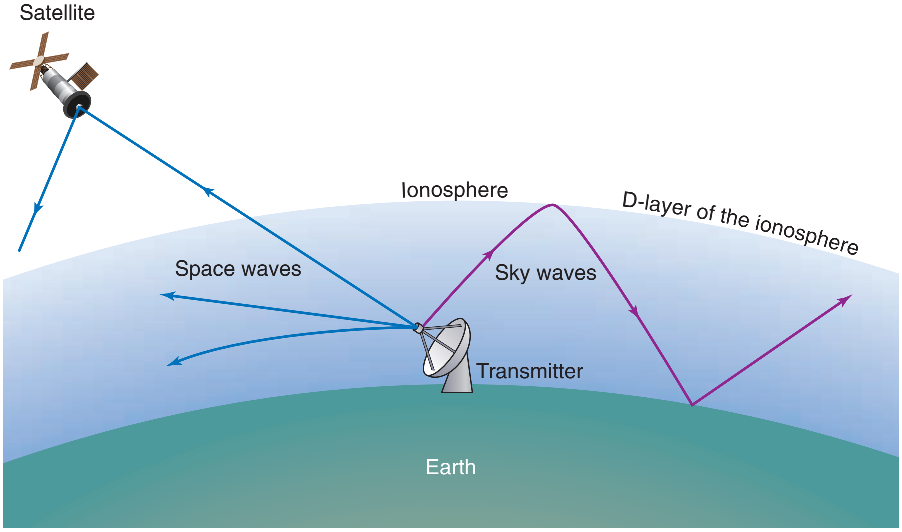


Figure 4.23 *Radio waves used in communication. Sky waves bounce off the ionosphere and the surface of the Earth and hence, bounce around the globe.*

4.3 REFRACTION

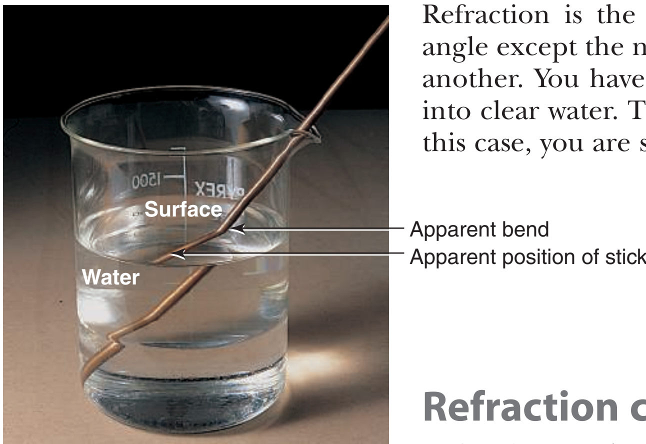


Figure 4.24 *An example of light waves refracting. This leads to an apparent depth in the water that is always less than the true depth.*

Refraction is the phenomenon where waves that are incident on any angle except the normal bend as they pass from one medium or depth to another. You have seen the bending effect when you put a straight stick into clear water. This apparent bending effect is shown in figure 4.24. In this case, you are seeing the refraction of light rays.

Refraction can be seen in water waves

Refraction can be seen in surface water waves where the water changes depth. With water waves it is possible to see the waves change in velocity where a change in depth occurs (see figure 4.25).

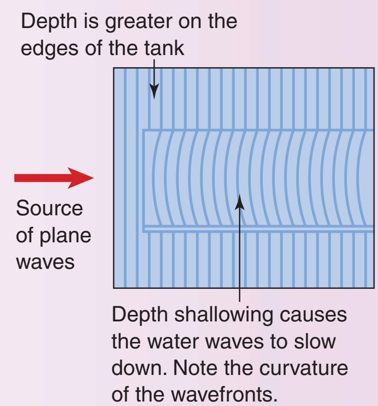


Figure 4.25 *Plane water waves in a ripple tank bunch up in the shallower water. Note that the frequency or number of waves doesn't change when the waves bunch up. It is only the wavelength and speed of the waves moving forward that change.*

4.3

Refraction in water waves

If the incident water waves strike the shallow-water region at an angle, the same slowing effect on the waves occurs, but now the wavefronts have bent. The waves in the shallow water have a shorter wavelength compared to those in deeper water but still maintain a constant frequency (see figure 4.26).

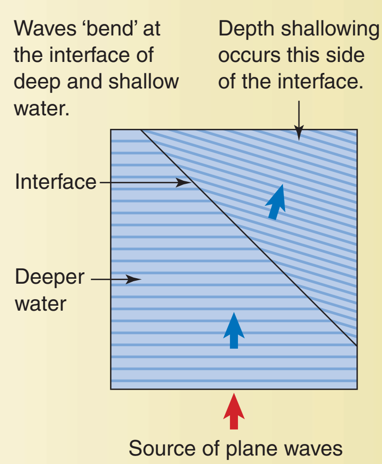


Figure 4.26 *Plane waves travelling from deep to shallow water. As each wavefront encounters the shallow water it slows down, creating a bend in the wavefront. This bending of the wavefront is refraction.*

The refraction of water waves is characteristic of the behaviour of all waves. The difference in water depths has the same effect as a change in the medium. A change in medium also affects the wave because each type of wave has a fixed velocity in any given medium. The wave velocity changes when a wave moves across an interface from one medium into another. The frequency of the wave does not change. Therefore, from the wave equation $f = \frac{v}{\lambda}$ the wavelength, λ, must change.

The apparent bending of the object at the interface between air and water, as shown in figure 4.24 on the previous page, suggests that, in a similar way to water waves entering a shallow medium and slowing down, light entering water is bent. Therefore, it must also have slowed down.

Consider a situation where plane waves move from a medium where they travel at high velocity into a medium where they travel at a lower velocity. The waves strike the interface at 90°, as shown in figure 4.27.

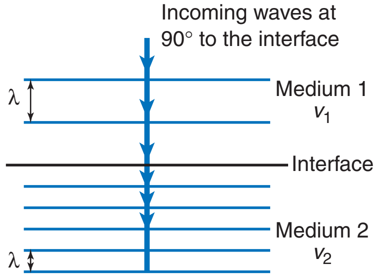


Figure 4.27 *Wavefronts passing an interface between two different media. The wavelength of the wave is larger in medium 1 than in medium 2.*

Notice how the wavelength decreases in medium 2. This implies that the velocity in medium 1 is greater than the velocity in medium 2 as the frequency of the wave remains constant.

If wavefronts strike a boundary between two media at an angle other than 90°, a change in wave direction occurs at the interface, along with a decrease in wave speed and wavelength. This is shown in figure 4.28.

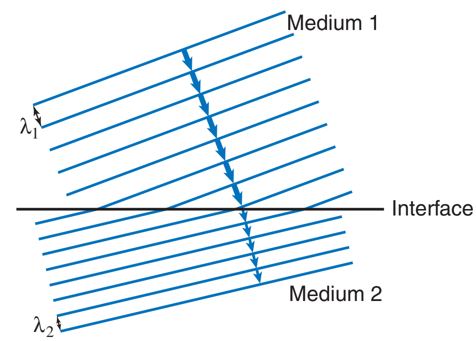


Figure 4.28 *In this case, the wavefronts slow down as they strike the interface. In this figure the wave velocity in medium 1 is greater than the wave velocity in medium 2. The wave direction bends towards the normal to the interface.*

To make it easier to study refraction, ray diagrams are used just as they are when dealing with reflection. Diagrams using wavefronts are difficult to analyse, so rays are used to represent the wavefronts instead.

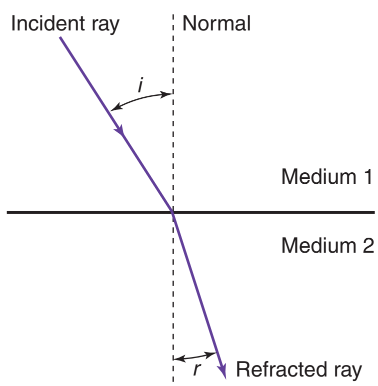


Figure 4.29 *Refraction of a ray that passes into a medium where its velocity is lower. Note that the bending of the ray representing the wavefront is towards the normal and occurs at the interface where the medium changes.*

When a wave moves from one medium to another where its speed is lower, the ray bends towards the normal. When a wave moves from one medium to another where its speed is higher, the ray bends away from the normal. Using rays you can easily measure an angle of incidence and refraction compared to a line drawn normal to the interface at the point of ray incidence. The rays are drawn perpendicular to wavefronts and the angle of incidence and the angle of refraction are both measured from the normal at the point of incidence (see figures 4.29 and 4.30).

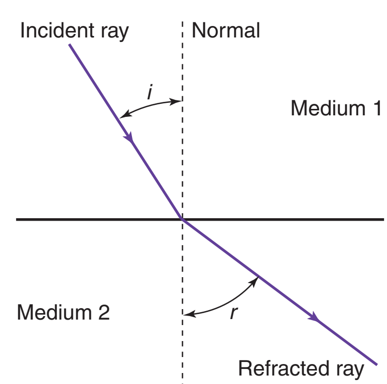


Figure 4.30 *Refraction of waves away from the normal. This occurs when the speed of the wave is slower in medium 1 than medium 2. Note the angle of incidence,* **i**, *is less than the angle of refraction,* **r.**

4.4 REFRACTIVE INDEX AND SNELL'S LAW

The relationship between speed, wavelength and angles of incidence and refraction was determined experimentally in 1621 by Dutch mathematician and physicist Willebrord Snell (1591–1626), and today it is known as Snell's Law. Snell's Law can be expressed mathematically, as:

$$\frac{v_1}{v_2} = \frac{\sin i}{\sin r}.$$

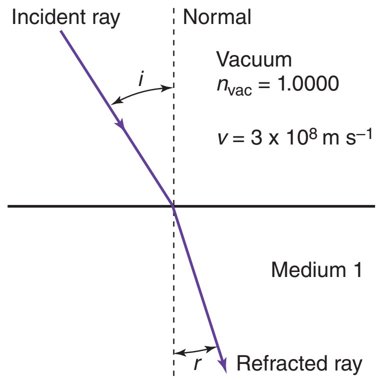


Figure 4.31 ***The absolute refractive index of a substance is determined by comparing the speed of the electromagnetic wave in that substance to the speed of the same electromagnetic wave in a vacuum such as space. Because the refractive index is derived by dividing a speed by a speed, the unit has no dimensions.***

PRACTICAL ACTIVITIES

4.4

Determining refractive index of glass

eBook*plus*

eLesson:
Refraction and Snell's Law
eles-0037
Interactivity:
Refraction and Snell's Law
int-0056

If a ray passes from a vacuum to another material of fixed composition and density, the degree of bending that occurs at the interface between the vacuum and the material is a constant. This constant is given the symbol n and is known as the absolute refractive index. This means that all refractive indices are measured with respect to a vacuum. A vacuum by definition has an absolute refractive index of 1.0000 for the electromagnetic spectrum. The absolute refractive index of any transparent material is a measure, or ratio, of how much an electromagnetic wave slows down at the interface between a vacuum and that material (see figure 4.31). That is, the absolute refractive index of some material, n_1, is: $n_1 = \dfrac{v_{vac}}{v_1}$.

Note that a refractive index is a comparative measure of the velocities of electromagnetic radiation in a vacuum and in the material. That is, in this case it is referring to the velocity of electromagnetic radiation relative to the velocity of electromagnetic radiation in a vacuum.

Table 4.1 lists the absolute refractive indices for some common materials.

Table 4.1 **Refractive indices of selected materials**

SUBSTANCE	VELOCITY OF LIGHT IN MEDIUM (m s^{-1})	ABSOLUTE REFRACTIVE INDEX OF MEDIUM
Vacuum	3×10^8	1.000 000
Air	2.999×10^8	1.000 28
Water	2.26×10^8	1.33
Crown glass	1.97×10^8	1.52
Denser crown glass	1.92×10^8	1.56
Flint glass	1.86×10^8	1.61
Denser flint glass	1.72×10^8	1.74
Perspex	2×10^8	1.46
Diamond	1.24×10^8	2.42

Notice that the absolute refractive index for air is 1.000 28. This is very close to the refractive index for a vacuum, hence you can use the approximation that the refractive index of air is 1.0 for most purposes and compare the refractive index of other substances to that of air. In other words, assume $n_{air} = 1.0$.

Snell's Law applies equally to waves slowing down and speeding up as they move across the interface between one medium and another. As with water waves at an interface, the frequency of the waves does not change as they speed up or slow down, so it is the wavelength of the wave that changes. This is expressed in the Snell's Law relationship: $\dfrac{v_1}{v_2} = \dfrac{\sin i}{\sin r} = \dfrac{\lambda_1}{\lambda_2}$.

The refractive index is useful for determining what will happen to electromagnetic waves that pass across an interface between transparent materials. The absolute refractive indices can be used directly to determine a number of factors. This comes about because $\dfrac{v_1}{v_2} = \dfrac{\sin i}{\sin r} = \dfrac{\lambda_1}{\lambda_2} = \dfrac{n_2}{n_1}$.

SAMPLE PROBLEM 4.1

Refraction and the speed of light in water

Light enters perspex, from air, at an angle of 40° (see figure 4.32). Using the information shown in the diagram and from table 4.1, determine:
(a) the angle of refraction
(b) the speed of light in the perspex.

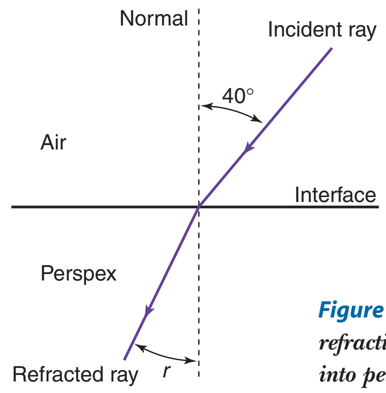


Figure 4.32 *Ray diagram showing refraction of light as it passes from air into perspex*

SOLUTION

(a) By Snell's Law:

$$\frac{v_1}{v_2} = \frac{\sin\ i}{\sin\ r} = \frac{\lambda_1}{\lambda_2} = \frac{n_2}{n_1} = \text{constant.}$$

Since the light here is going from air to perspex, then air is medium 1 and perspex is medium 2:

$$\frac{v_a}{v_p} = \frac{\sin\ i}{\sin\ r} = \frac{\lambda_a}{\lambda_p} = \frac{n_p}{n_a} = \text{constant.}$$

Using:

$i = 40°$
$n_p = 1.46 \leftarrow$ (from table 4.1)
$n_a = 1.00.$

Then:

$$\frac{\sin\ i}{\sin\ r} = \frac{n_p}{n_a}$$

$$\begin{aligned}\sin\ r &= \frac{n_a \sin\ i}{n_p}\\ &= \frac{1.00 \times \sin 40°}{1.46}\\ &= 0.4403\\ r &= \sin^{-1} 0.4403\\ &= 26.12°.\end{aligned}$$

The angle of refraction is 26.1°.

(b) The ratio of the velocities can be determined directly from the ratio of the absolute refractive indices of air and perspex.

$$\frac{v_a}{v_p} = \frac{n_p}{n_a}$$

$$\begin{aligned}v_p &= \frac{n_a \times v_a}{n_p}\\ &= \frac{1.00 \times 3.0 \times 10^8}{1.46}\\ &= 2.055 \times 10^8\ \text{m s}^{-1}\end{aligned}$$

PHYSICS FACT

The different parts of the eye have different refractive indices. The cornea has a refractive index of 1.37, the aqueous humour 1.33, the lens cover 1.38, the lens centre 1.41 and the vitreous humour 1.33. Light rays entering the eye to produce an image on the retina form an upside-down image and must pass through materials with all of these refractive-index changes. It is the sum of all of these changes in refraction that produces what is seen. Because of its multiple-refractive-index lens system and adjustable iris-aperture system, the eye has no problems with chromatic or spherical aberration.

The most significant interface, as far as focusing is concerned, is not the lens but rather the cornea because that is the interface where the greatest difference in refractive indices occurs (1 in air to 1.37 in the cornea). That is why laser surgery involving reshaping of the cornea can be used to correct many sight defects.

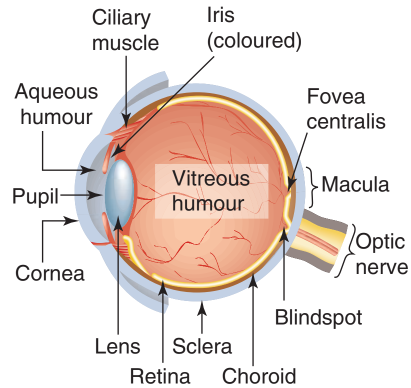


Figure 4.33 *A cross-section of the human eye*

4.5 TOTAL INTERNAL REFLECTION

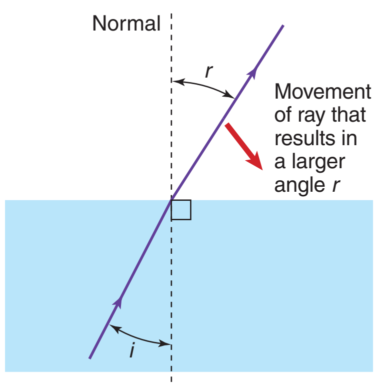


Figure 4.34 *A ray moving from a high-refractive-index medium such as glass to a lower refractive-index medium such as air*

You may have observed that when you look through a fish tank at a certain angle you see a reflection rather than being able to see through the clear water. A similar thing happens if you look at the face of a water-resistant watch when it is underwater. You can see the time until you turn the angle of the watch to beyond some critical angle. At this point the face of the watch takes on a mirror-like appearance. Both of these examples are cases of total internal reflection.

Total internal reflection is a case of Snell's Law in operation. It may occur when a ray of light attempts to cross into a low-refractive-index medium from a high-refractive-index medium. The ray bends away from the normal in this circumstance.

Because a small change in the angle of incidence causes a larger change in the angle of refraction when the situation shown in figure 4.34 occurs, it is possible for an angle of incidence to reach a critical angle where the ray can't exit the higher refractive-index material. The ray is refracted so much that it is bent to 90° from the normal at the interface (see figure 4.35).

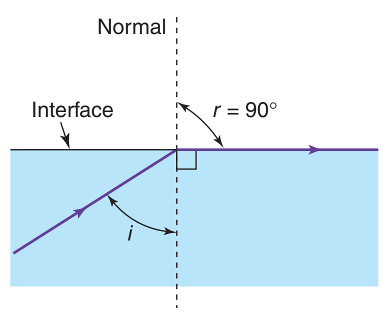


Figure 4.35 *A ray incident at a change in medium from higher refractive-index material to lower refractive-index material under the conditions shown cannot escape the high-refractive-index material and in such a case is forced to travel along the interface.*

The critical angle is the angle where total internal reflection prevents the ray from escaping from a higher optical-density medium to a lower optical-density medium.

The angle of incidence where the ray is just trapped in the higher refractive-index material is the **critical angle**.

If the critical angle of a ray at the interface of two substances is exceeded, the interface will act as a mirror and total internal reflection of light rays occurs. The ray then obeys the Law of Reflection at the interface between the materials. The ray is trapped internally within the denser material, as shown in figure 4.36.

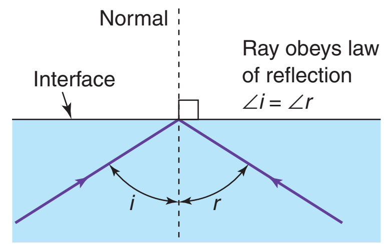


Figure 4.36 *A ray undergoing total internal reflection at the interface of a high refractive-index material and lower refractive-index material is trapped and reflected in the denser material.*

Calculating the critical angle

You can find the critical angle of refraction at a boundary between two media using Snell's Law:

$$\frac{v_1}{v_2} = \frac{\sin\ i}{\sin\ r} = \frac{\lambda_1}{\lambda_2} = \frac{n_2}{n_1}$$

where
1 and 2 subscripts = refractive index values for the first and second materials the light enters
i = angle of incidence
r = angle of refraction.

At the critical angle, the angle of refraction is 90°. Sin 90° is equal to 1 so the equation simplifies to:

$$\frac{v_1}{v_2} = \frac{\sin\ i}{1} = \frac{\lambda_1}{\lambda_2} = \frac{n_2}{n_1}.$$

Thus $\sin i_c = \frac{n_2}{n_1}$ where i_c is the critical angle.

Note that $n_2 < n_1$ for light to bend away from the normal. This simple calculation will give the critical angle.

Refraction leading to total internal reflection is shown in figure 4.37. Total internal reflection of the incident ray at the interface occurs when the angle of incidence is greater than the critical angle.

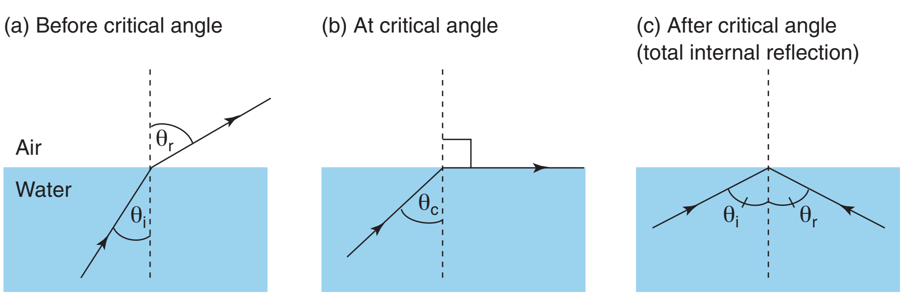


Figure 4.37 *The three possible results of a ray travelling from a high-refractive-index material, such as water, to an interface with a lower refractive-index material such as air*

4.6 OPTICAL FIBRES AND TOTAL INTERNAL REFLECTION

Total internal reflection has a number of practical uses. One of those is in optical fibres. This application of total internal reflection of electromagnetic waves has led to optical fibres becoming a major data carrier in telecommunications.

Optical fibres are made from thin, cylindrical strands of ultra-high-purity glass (see the photo in figure 4.1, page 46). These optical fibres are made so that they have a central, high-refractive-index region called a core. Their outer region, called the cladding, is made from a lower refractive-index glass. After the electromagnetic radiation enters the optical fibre, it is totally internally reflected at the interface between the higher refractive-index core and the lower refractive-index cladding. This means that rather than escaping through the surface of the optical fibre, the light is trapped internally, and continually moves forward through the optical fibre. This is shown in figure 4.38. To allow this transfer of trapped light to be faster and more efficient, the path-length travelled by the narrow beam of light is reduced. This is done by making the diameter of the core around 10 μm. This results in only those rays travelling parallel or almost parallel to the axis of the fibre being successfully guided along the fibre.

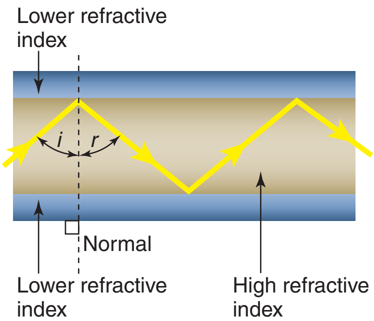


Figure 4.38 ***The cross-section of an optical fibre. The light waves transmitted by an optical fibre are reflected off the boundary between the high- and low-refractive-index material. The smaller the refractive index of the cladding (compared to the refractive index of the core), the smaller the critical angle. This allows light to be totally internally reflected much more easily. A small critical angle is vital for total internal reflection to occur easily. Note that the internally-reflected ray obeys the Law of Reflection.***

Optical fibres are used in a number of situations, including:

- communication for carrying signals precisely, and at the speed of light. This is faster than energy transmission by electrons in electric signals.
- medicine. Operating doctors view sites such as the intestines, previously inaccessible without invasive surgery, by using optical fibres in instruments known as endoscopes.

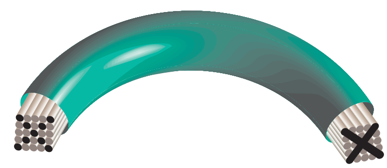


Figure 4.39 ***A bundle of optical fibres with each fibre carrying its individual part of the image. The arrangement of the bundled fibres must remain constant to reproduce an image of the object.***

Optical fibres are particularly suited to their role because they allow the transmission of light to or from locations where straight line transmission of the light would not be possible. They are also quite flexible which allows electromagnetic radiation and, in particular, visible and infra-red radiation to be reflected easily and precisely around corners without the need for any physical reflective device such as a silvered mirror.

4.7 DIGITAL COMMUNICATION SYSTEMS

Digital communication systems are based on signals that have two values, on or off (1s and 0s).

The trend in all communication signalling is towards **digital** transmission and storage. Almost all long-range communication systems now in use transmit the signal as a digital pulse. Digital signals are based on a binary code of 1s and 0s.

There are excellent reasons for favouring digital communication transmissions. These reasons include enhanced security as well as the ability to preserve the signal and still be able to read it after it has suffered more interference and superposition than is the case with **analogue** signals.

Analogue signals send information over a continuous range of values.

If a digital signal is subjected to interference or degraded due to attenuation occurring over a long distance then the receiving instrument can still generally tell whether the pulse represents a 1 or a 0. The message can then be decoded or if necessary, be cleaned up and re-transmitted by a repeater station without loss of information from the original transmission.

If an analogue signal has information added to it (due to interference), the actual shape of the wave, and hence the data, has been changed. Such changes are much more difficult to remove than the additions to digital signals. Also, if the transmission is to be re-transmitted, the noise cannot be separated from the original signal and so is amplified and forwarded on with the transmission. This can lead to rapid degradation of the original signal after only a few re-transmissions. However, a re-transmitted digital signal from a repeater station is identical to the original transmission.

The digital revolution

Switching from analogue to digital communication devices came about after the invention of a system to transport digital signals and the devices to encode and read digital information. Those two devices were the optical fibre and the laser.

In 1970, Corning Glass Works in the United States produced the first optical fibre that was suitable for long-range communication. Optical-fibre communication systems for transmitting data or sound use special lasers to transmit messages that are encoded as pulses of light. They can transmit more information than traditional copper-wire analogue systems which send information encoded as electrical pulses through large cables.

The data sent along optical-fibre cables is always a digitised signal. Mobile phones use a digital signal for data and voice transmission, television is now broadcast as a digital signal, and the internet works because of digital transmission along data and phone lines.

How does digital work?

All digital signals work in the same way. A sound, picture or data is initially generated as an analogue signal that is frequency- or amplitude-modulated. The digitising process is applied to the analogue signal. This process is called quantisation. In the quantisation process, the analogue wave shape is converted into a code that represents a set of numbers. These numbers can be represented in a binary code of 1s and 0s. In 8-bit processing, the signal is represented as a series of numbers from 0 to 7 that can be represented as binary digits using some combination of three 0s or 1s. The amplitude of an analogue wave at a particular time interval is scaled and converted to a number on a scale from 0 to 7 for 8-bit processing. This is called sampling. The sampled information is converted by a digital-encoding device to a sequence of a combination of three 0s and 1s. Sampling is done many times per second as the waveform of the analogue signal is encoded. The digital signal that represents the analogue wave can then be transmitted as a binary sequence of all the samples in sequential order. A decoding device can then reconstruct the amplitude of the analogue wave at that point in time in a sequence.

To increase the accuracy of the wave reconstruction to the original analogue signal, two things can be done. More samples can be taken at closer time intervals. The number of divisions on the vertical scale can also be increased by making the processing 16-bit, 32-bit or 64-bit.

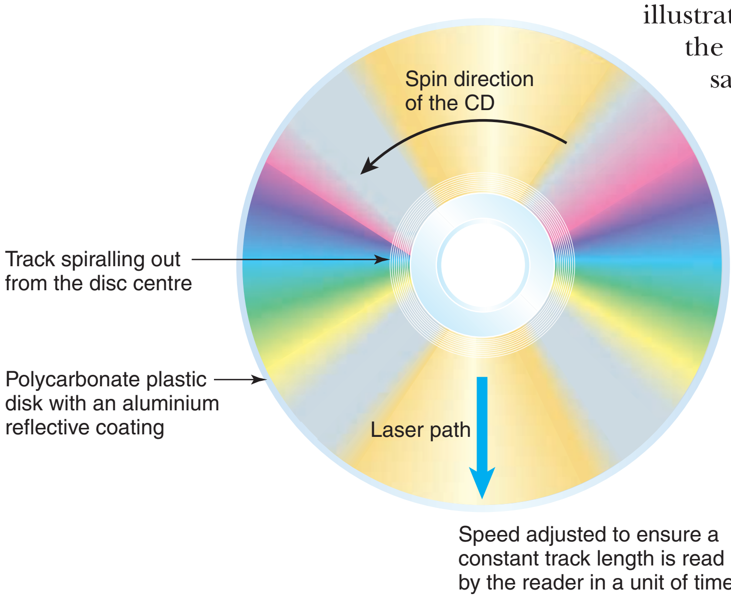
The invention and commercialisation of the laser has meant a digital revolution in sight and sound transmission and storage. The revolution has involved the development of CD and DVD technology.

Disc-shaped digital storage devices called CDs and DVDs store and allow for the reproduction of the highest possible quality sound and pictures. Because the original analogue signal is not retained on these devices which instead use a code of the signal, the reproduction is perfect.

The compact disc (CD)

The compact disc is a plastic, metal-coated disc that stores information digitally. The information (that can be converted to pictures or sound) is stored as a series of pits, representing 1s, on a spiralling track. This is illustrated in figure 4.40. The value of each length of the spiralling track or each digit 1 or 0 is always the same, so the lack of a pit burned into a track represents a 0. The principle of storing and reading the data is the same whether the device is converting the signal from audio, data or vision transmissions.

Recording an audio CD involves a microphone that translates sound waves into analogue (wave-like) electric signals. An encoder then divides the wave signal into 44 100 segments for each second of sound. Each of the 44 100 segments (representing a single second of sound) is then converted into a digital code according to the amplitude of the analogue signal at that time. During disc writing, a pulsing laser uses this code as an on/off template as it cuts a spiral track of microscopic pits into the surface of the disc. The pits are the digital code representing a 1 or a 0. When playing an audio CD, the disc spins in a CD player while a laser beam shines on the pit spiral. The metal coating on the CD reflects the light. The intensity of the reflected light changes as the beam enters and leaves the pits and it is the difference in this reflected beam's intensity that translates into an electrical signal. Decoding circuits use this electrical signal to reproduce the original sound in loudspeakers or headphones.



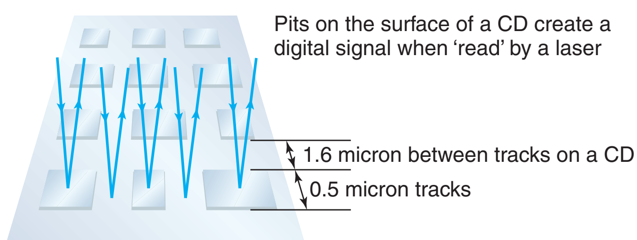


Figure 4.40 *A compact disc*

Features of a DVD

A digital versatile disc (DVD) is the same size as a compact disc but has a much greater storage capacity. The increased storage capacity comes about because the track on a DVD is much longer than that on a CD, the pits are smaller and the spirals of the track are much closer together. Figure 4.41 outlines how a CD or DVD works.

The revolution in digital storage media has led to development of new technologies such as the Blu-ray disc that can store 25GB on a single-layer disc and 50GB on a dual-layer disc. Blu-ray discs use a blue violet laser (405 nm) that can focus with greater precision, enabling more data to be packed onto the same-sized disc.

The spiral data tracks on a standard DVD are only 740 nm apart, providing a high-density data medium with a minimum pit length of 400 nm and width of 320 nm. In dual-layered discs the laser reads the first layer, then passes through a semi-transparent gold layer to read the second layer that is slightly deeper in the plastic disc. The total track length of a double-sided, dual-layered DVD is of the order of 50 km. Standard DVD uses a red laser operating at a wavelength of 650 nm.

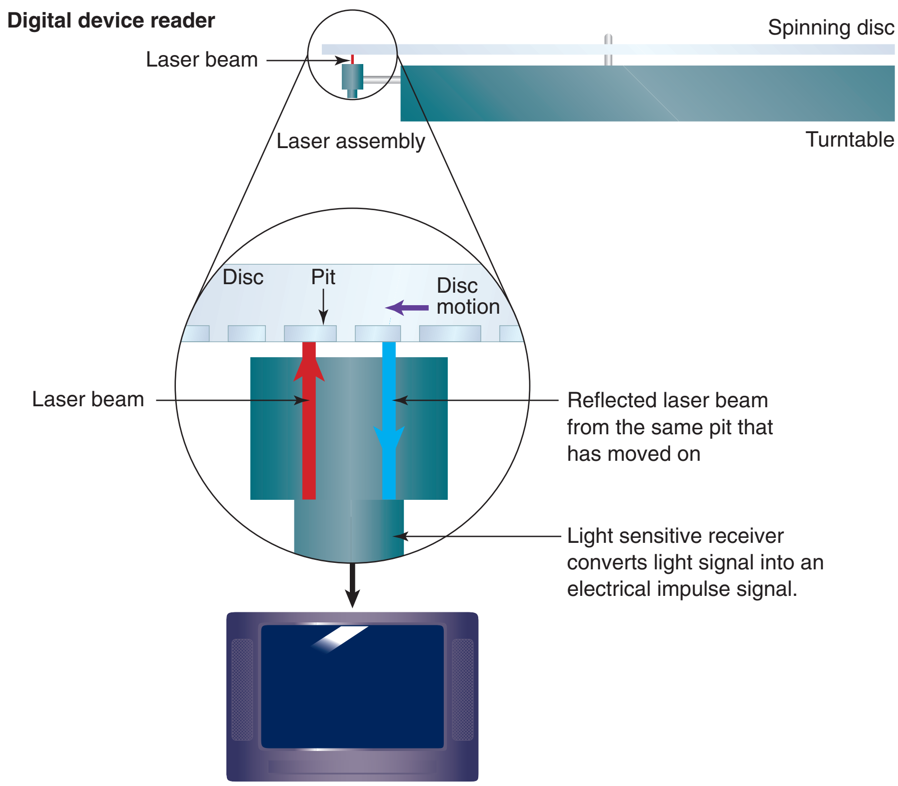


Figure 4.41 *How a CD or DVD reads the signal. A disc player converts the pictures and sounds recorded in digital form on a disc into a visible or audible signal. The laser inside the reader aims a very thin beam of light at the spinning disc, precisely following the spiralling pit track out from the centre of the disc. Another device then reads the pattern of light reflected off the disc and converts the information into a useful form that is read and processed by the decoding machine.*

PHYSICS FACT

The trend in Australia is towards the broadcast of digital television. Analogue television broadcast in Australia is planned to be switched off in 2013. Digital broadcast signals are transmitted as a series of 1s and 0s. Viewing digital television means you need to have a digital receiver associated with your display panel.

Digital television has several advantages over analogue television:

- Digital channels can take up less bandwidth. Broadcasters may choose to send a high definition image (HDTV) on a single channel; or in the same spectrum space, send multiple lower standard definition (SDTV) quality images on multiple channels. Broadcasters can provide in the same amount of spectrum space more digital channels, provide the best possible image and sound (HDTV), or provide other non-television services such as electronic program guides, subtitles and additional languages.
- The bandwidth needs of broadcasts are variable for digital TV. A talking head presenting the news has much less changing screen information than an action packed sporting event such as a football game. Digital television allows broadcasters to send broadcast signals that are continuously variable in terms of the amount of signal required to produce the best image.
- Sound on digital TV is CD quality.
- Digital television is broadcast in 16:9 format and is picture perfect for as long as it is received. Analogue signals degrade as you go further from the broadcast antennae and are subject to interference. As long as the digital signal is received the image is perfect. When failure to receive the signal occurs the image pixelates or fails completely.

Telecommunication transmissions

Telecommunication transmissions are either analogue or digital. Analogue transmission uses signals that are exact reproductions of the electrical signal produced in the transducing device (by the sound or image being transmitted). An analogue telephone system transmits an electric current that copies the pattern of electrical signals that sound waves of the speaker's voice produce in the microphone. This current, after passing over the copper-wire network, is converted back to sound waves by the speaker in the receiving telephone.

Digital transmissions are different from analogue transmissions because the signals are converted into a code of 0s and 1s before transmission. In the case of a transmission over an optical-fibre cable, the code is like an on–off flashing of a light. A laser produces the light that then travels through thin strands of glass called optical fibres. Transmitting a telephone conversation requires the light in the system to flash on and off about 64 000 times per second. Although this may seem very rapid, it is common for optical fibres to be able to carry about 6000 conversations at the same time. This requires the laser to flash on and off about 41 million times per second.

Global positioning system (GPS)

Figure 4.42 ***A hand-held GPS***

The global positioning system (GPS) is made up of a set of satellites orbiting the Earth in precise locations at elevations of about 17 600 km. The location of these satellites is such that the entire surface of the globe has a direct line-of-sight connection between at least some of the satellites. To locate your position on the surface of the Earth you must have a GPS receiver and be in line-of-sight contact with at least three satellites. This is necessary for the GPS receiver because it only detects line-of-sight transmissions from the satellites.

The satellites act in conjunction with a series of ground stations that are in constant communication with the satellites by radio. These radio communications tell the satellites exactly where they are with respect to the surface of the Earth at any time. The whole system works on the principle of knowing the exact location of the satellites at a precisely known time. Each satellite has an on-board atomic clock that gives it a precise time base. Each satellite is constantly broadcasting a radio signal that contains information about the time that the signal was sent and from which satellite the signal was sent. Software corrects for any delays the radio signal experiences as it travels through the atmosphere on its way to GPS receivers. The requirement for this correction means that the more satellites your GPS receiver can receive radio signals from at the one time, the more accurately it can determine your position.

Because of the satellites' high orbit and the curvature of the Earth, a GPS receiver may see up to nine satellites at any time. A minimum of three satellites must be seen to locate a position on the Earth's surface. A minimum of four satellites can also give information about the altitude of the GPS receiver. The GPS can locate a position accurately because the speed of transmission of the radio signal from the GPS is known. Hence, the distance of each satellite from the GPS receiver on an imaginary sphere drawn on the surface of the Earth can be very accurately calculated. The position of the GPS receiver is determined by the intersection of the different spheres of possible position suggested by each different GPS satellite (triangulation), as shown in figure 4.43.

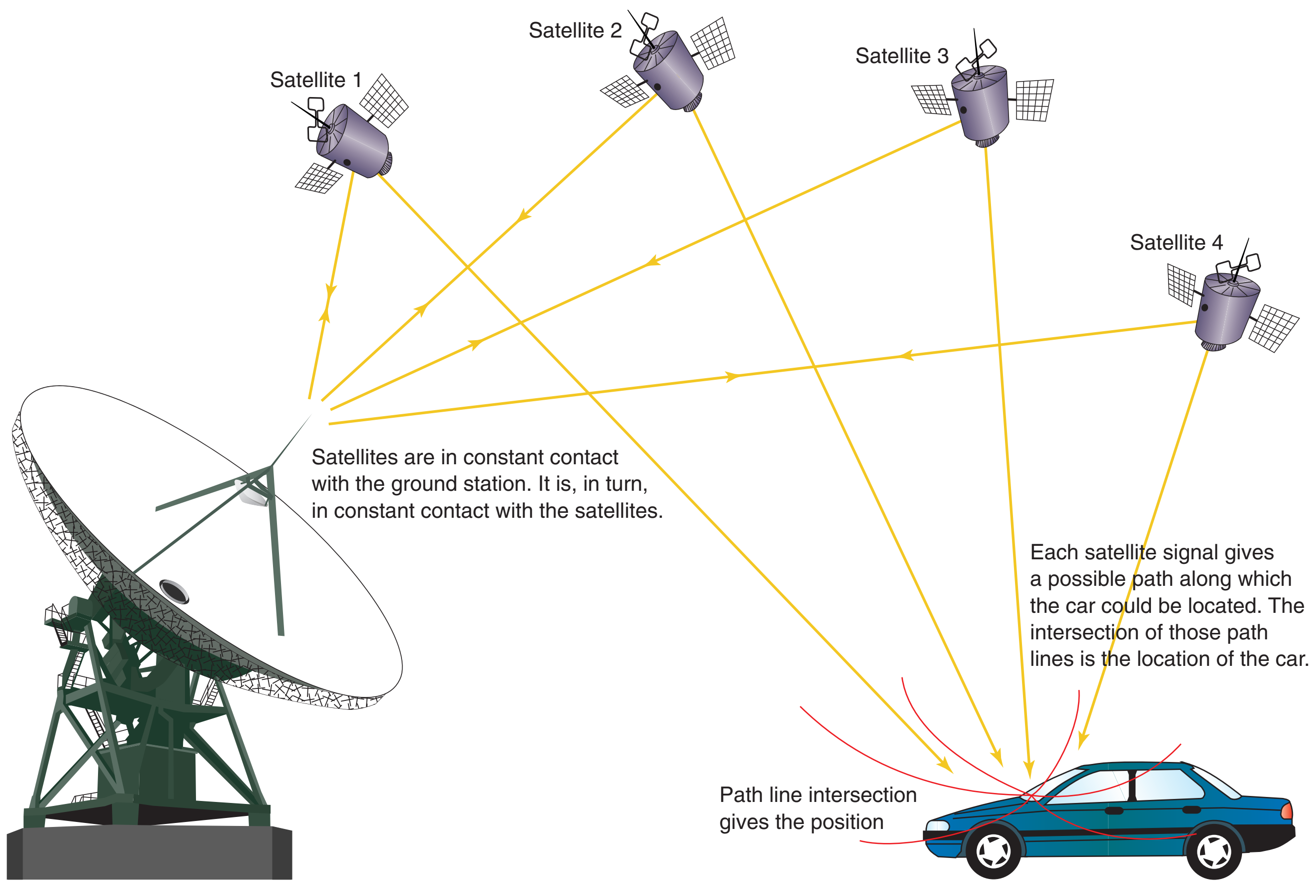


Figure 4.43 *The GPS in operation*

Combining small hand-held mobile computing devices, clever software and detailed maps with voice simulation has resulted in one very clever GPS advance; the in-car navigation system.

SUMMARY

- The Law of Reflection states that the angle of reflection is equal to the angle of incidence.
- Reflective regular surfaces can be diverging, converging or plane.
- Light and other electromagnetic waves are used to send digital signals along fibre-optic cables.
- Light waves are used to read digital signals from digital recording media such as CDs and DVDs.
- Analogue signals must be modulated to carry signal information. This can mean the amplitude is modulated or it can mean the frequency is modulated.
- Refraction is the bending of waves at the interface of media transparent to those waves.
- Refraction results from a difference in the velocity of the wave in the different media it encounters. Where a wave is slowed down across an interface, it bends towards the normal to the interface. Where a wave speeds up across an interface, it bends away from the normal.
- The absolute refractive index of light is the speed of light in a medium compared to the speed of light in a vacuum.
- Snell's Law states that $\frac{v_1}{v_2} = \frac{\sin i}{\sin r}$ where v_1 is the velocity of the light in the incident medium and v_2 is the velocity of light in the refractive medium.
- The critical angle is the angle of incidence which gives an angle of refraction of 90°.
- If the angle of incidence is such that it exceeds the angle where $r = 90°$ when the ray reaches a high-refractive-index/low-refractive-index interface, then total internal reflection will result and the ray will not be able to escape the higher refractive-index material.
- Optical fibres work on the principle of total internal reflection and consist of a high-refractive-index core surrounded by a lower refractive-index cladding that traps the light in the optical fibre.

QUESTIONS

1. Describe the Law of Reflection.
2. A ray of light from a laser hits the surface of a plane mirror at an angle of 40°. Calculate at what angle to the surface the laser beam will be reflected.
3. Describe the features of a reflection of an object that is reflected from a plane mirror, compared to the original object.
4. Identify a way in which each of the following electromagnetic wave types is used in communication and information transfer:
   (a) light
   (b) radio waves.
5. Describe how reflection from each of these types of surfaces is used as an aid to communicate:
   (a) plane surfaces
   (b) concave surfaces
   (c) convex surfaces
   (d) the ionosphere.
6. Define the term refraction.
7. Snell's Law is represented in symbols as:

   $$\frac{v_1}{v_2} = \frac{\sin i}{\sin r} = \frac{\lambda_1}{\lambda_2} = \frac{n_2}{n_1} = \text{constant.}$$

   Explain this fully in words.
8. When a wavefront strikes the interface between two different media at an angle, the wavelength changes. Describe what happens to the frequency of these waves and how this explains the wave bending towards or away from the normal to the interface at the point of intersection.
9. Describe the conditions necessary for total internal reflection.
10. Outline how total internal reflection is used in optical-fibre technology.
11. Present a ray diagram to scale that will enable you to describe in full the image of a 4 cm tall object located 15 cm from the pole of a concave mirror.
12. Present a ray diagram to locate the image of an object 5 cm high located 8 cm from a 5-cm focal-length mirror. Describe the nature of that image in full.

13. A girl looking into a concave mirror notices her image is greatly magnified and the correct way up when she is at position (a) then her image disappears completely as she walks away from the mirror for a split second at position (b) then is inverted and magnified at position (c) before becoming inverted and diminished at position (d). Present a sketch of the mirror showing the pole, focus, centre of curvature and location of positions (a) to (d).

14. Calculate how fast light travels in a material with a refractive index of 1.47.

15. A ray of light is incident on an air/glass interface at an angle of 30° to the normal. Calculate the angle of refraction of the ray in the slab The refractive index of glass is 1.50; the refractive index of air is 1.00.

16. Calculate the critical angle at which a ray of light must hit a glass/air interface to be totally internally reflected. The refractive index of glass is 1.50; the refractive index of air is 1.00.

17. Identify a number of communications that are made using a digital system during transmission.

18. Describe the technological developments that were necessary before the GPS became a reality.

19. Explain why the corruption of a digital signal is less likely than corruption of an analogue signal (in relation to the extent of change in the received signal from the original signal transmitted).

20. Describe the technological developments that were necessary before digital storage on media such as CDs and DVDs was possible.

4.1 USING RAYS TO LOCATE AN IMAGE FROM A PLANE MIRROR

Aim

To locate the position of an image in a plane mirror using a geometrical construction and the Law of Reflection

Apparatus

small plane mirror on a stand (if the mirror is not front silvered you will have to make allowance for the protective glass when interpreting your findings)
four pins
newspaper
sheet of writing paper
ruler
pencil
protractor

Theory

The angle of incidence should equal the angle of reflection from a plane mirror.

Method

1. Use a pin located at a point, P, as a source of light rays. Stick the pin in the paper sheet on which you intend to draw your construction so that it cannot move. To make this simpler, do this activity on a newspaper with a clean sheet of paper overlain.
2. Find the position of its image in a plane mirror represented by the symbol P′ by drawing the three construction rays shown in figure 4.44. The first of these is a line drawn from the pin to the mirror to form a normal. It will be necessary to extend the normal line behind the mirror. Draw the second and third lines randomly but in the general direction (as shown) towards the mirror so they appear to originate at the pin and travel towards the mirror.
3. Move your line of vision so that it appears as though you are looking along one of the reflected randomly drawn lines towards the mirror and the pin. Draw in that line. These lines are easily seen as the reflected pencil line leading to the reflected image of the pin.
4. Repeat the procedure for the other randomly drawn reflected line.
5. Remove the mirror and extend the normal and the traced lines of the two randomly drawn line reflections behind where the mirror was located as shown. Where all three lines intersect behind the mirror is the apparent position of the image behind the mirror.

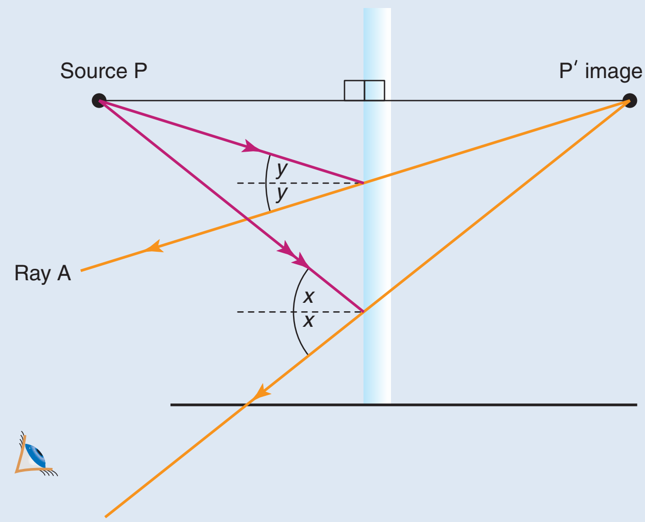


Figure 4.44 *Image formation in a plane mirror*

Analysis

The image in the mirror is not a real image. This means P′ is a virtual image. The rays of light producing the image do not really pass through or derive from the point P′, they only appear as if they do.

Questions

1. How do you know the image produced is not a real image?
2. How could you improve on this experiment?

4.2 REFLECTION IN CONCAVE AND CONVEX MIRRORS

Aim

To observe light reflected from a concave and a convex mirror

Apparatus

ray box and a power pack
a three-slit screen to enable you to produce three narrow beams of light
convex mirror (one is supplied with ray box kits)
concave mirror (one is supplied with ray box kits)

Theory

The law of reflection should be obeyed for each incident ray. The focus of the concave mirror can be found in front of the mirror and the focus is behind the mirror for the convex mirror.

Method

1. Place the three-slit screen into the slot of the ray box on the side that has the focusing lens. This should enable you to see three parallel beams of light.
2. Shine the beams of light along a sheet of flat white paper towards a *concave mirror* so that the beams strike the concave mirror in such a way that the centre beam hits the pole of the mirror and the two side beams are still reflected. If you do this successfully, the centre beam will be reflected straight back. Adjust the position of the mirror until it is reflected straight back. This is the principal axis of the mirror.
3. Draw what you see on the paper by tracing the front of the mirror and the beam paths with a ruler. You may find it easier to mark each beam's position (incident and reflected) with a pencil in two places and draw them in later with a ruler.
4. Arrange the three beams so that they now reflect off the *convex mirror* in a similar way to procedure 2. Again, aim the central beam at the pole of the mirror and it should be reflected straight back. Adjust the position of the mirror until it is reflected straight back. This is again the principal axis of the mirror.
5. Draw what you see by tracing the front of the mirror and the beam paths with a ruler. You may find it easier to mark each beam's position (incident and reflected) with a pencil in two places and draw them in later with a ruler.

Analysis

1. Was the Law of Reflection strictly obeyed for all incoming rays?
2. How could this experiment be improved?
3. Was there any evidence of spherical aberration? If so, what was it?

Question

What was the focal length of each mirror?

4.3 REFRACTION IN WATER WAVES

Aim

To observe refraction occurring in water waves

Apparatus

ripple tank or shallow cooking tray
rectangular slab of glass
ruler or oscillator to produce regular water waves

Method

1. Place a rectangular glass slab in the tray and add water until the slab is covered to a depth of around 1 mm. The slab should be positioned so that the waves will hit the slab square on.
2. Set up a source of plane water waves, such as those produced by a ruler vibrating back and forwards with a rapid, regular frequency, at one end of the tray and begin to generate water waves.
3. Note what happens to the plane (straight) waves as they travel from the deep water to the shallow water (1 mm over the slab).
4. Repeat steps 1 to 3 but have the waves hitting the slab on an angle by adjusting the position of the slab.

Analysis

What do you notice about the wavelength, frequency and speed of the waves?

Questions

1. What happens to the waves hitting the deep-water/shallow-water interface perpendicularly?
2. What happens to the waves hitting the deep-water/shallow-water interface at an angle?

4.4 DETERMINING REFRACTIVE INDEX OF GLASS

Aim

To observe, measure and graph the angle of incidence and refraction for light encountering a medium change at an air/glass interface and to use the measurements to determine the refractive index of glass

Apparatus

either an optics kit or ray boxes, or try the following method using pins:

four dressmakers pins
rectangular slab of glass
ruler and a protractor

Theory

The refractive index of a ray produced by a pair of pins can be traced as the path of four pins through a glass slab. Two pins can be used on each side of the slab of glass to define the straight ray path.

Method

1. Place a rectangular slab of glass on a blank sheet of paper sitting on a soft surface (that can be damaged by inserting pins) and trace around the glass slab. A newspaper makes a suitable surface into which to insert the pins.
2. Keep the slab in the same place over the trace mark and on one side of it push two pins into the page, as shown in figure 4.45.

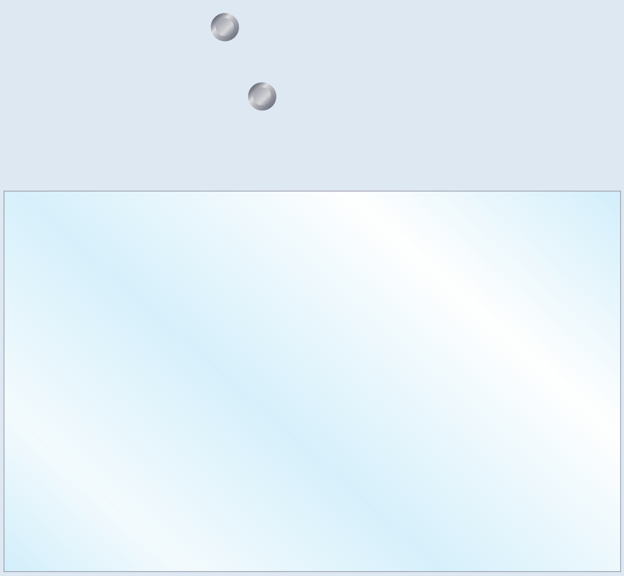


Figure 4.45 *The glass slab and pins inserted on one side*

3. Look through the slab from the opposite side to the pins. Move your head so your line of vision is such that the two pins pushed into the paper and underlying surface appear to line up directly behind one another through the glass slab.
4. When the pins on that side appear to be lined up, insert another two pins on the side of the slab from which you are looking so that all four pins appear to line up one behind the other as shown in figure 4.46.

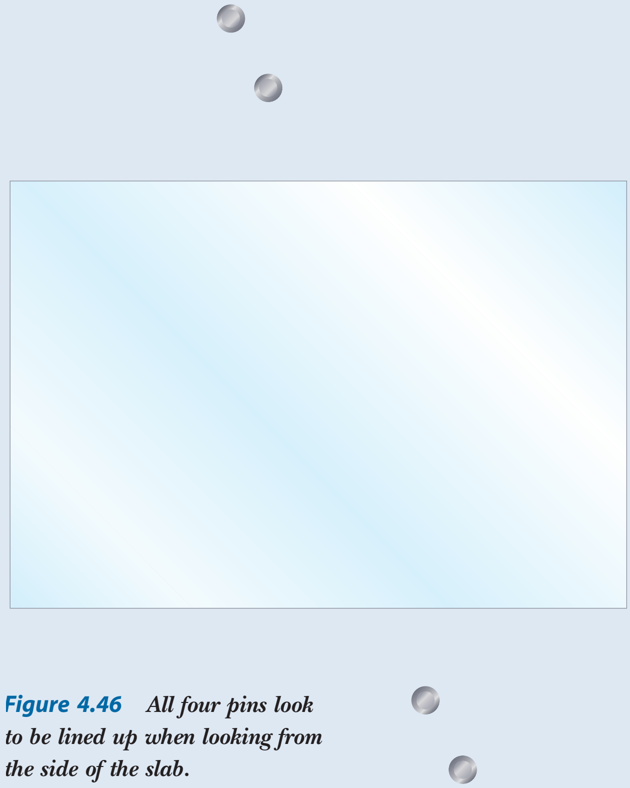


Figure 4.46 *All four pins look to be lined up when looking from the side of the slab.*

5. Remove the slab and pins, carefully noting where the pin holes are.
6. Rule a line through each set of two pin holes on each side of the slab so that the lines extend to the outline of the slab as shown in figure 4.47.

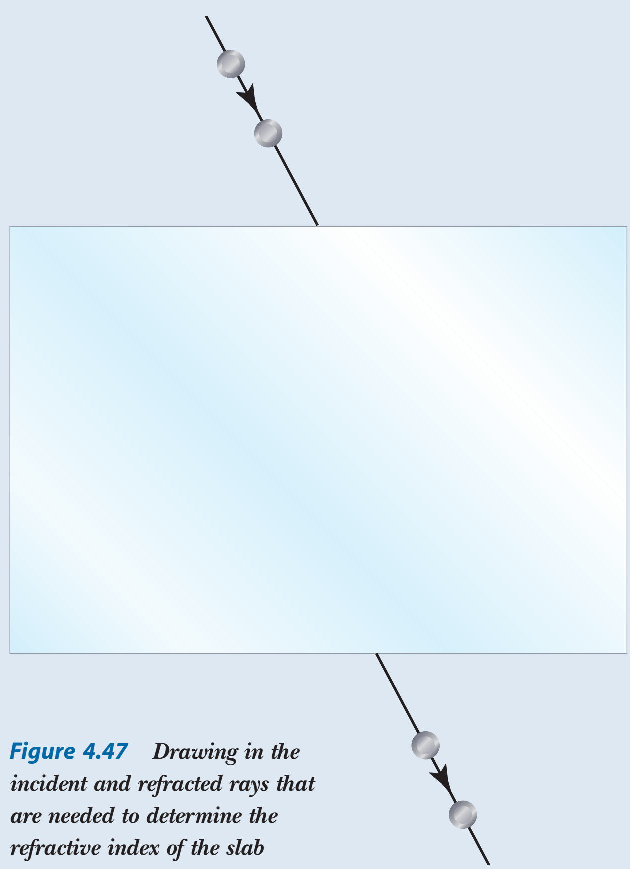


Figure 4.47 *Drawing in the incident and refracted rays that are needed to determine the refractive index of the slab*

7. Draw a line across the outline of the slab that joins the two lines as shown in figure 4.48.

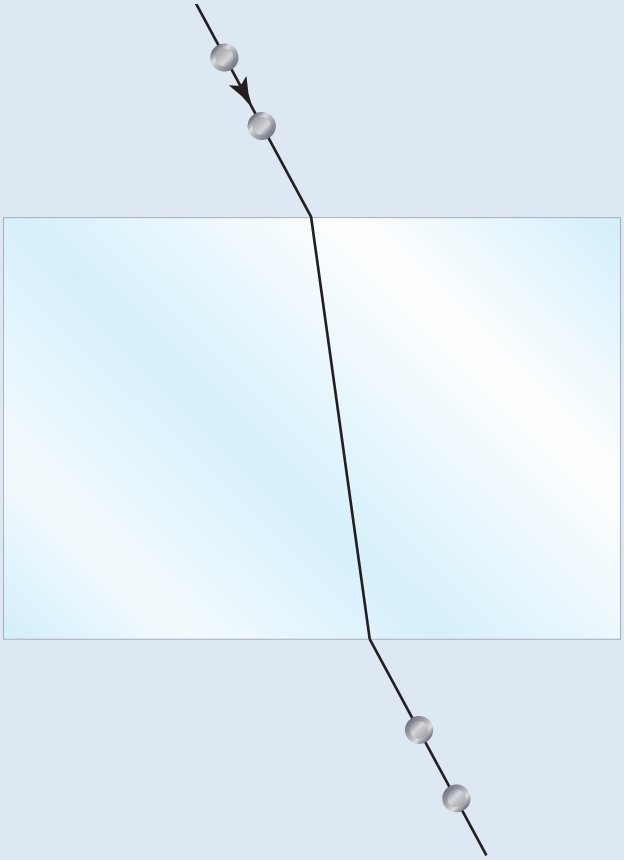


Figure 4.48 *The refraction of the light is apparent from the bend in the ray entering and leaving the slab. Note that the ray before entry and after exit from the slab are parallel.*

8. Draw in the normal to the slab edge where the two lines joining the pin holes meet the slab as shown in figure 4.49.
9. Mark in the angles of incidence and refraction as shown in figure 4.49 and measure the angles, i and r, using a protractor.
10. Calculate the refractive index of the slab using Snell's Law. Assume that the refractive index of air is 1.
11. Repeat the experiment using three different angles of incidence by changing the position of the two pins initially inserted and repeating the procedure.

Analysis

1. Calculate an average value for the refractive index of the glass slab based on your experimental work.
2. Compare the average value of the refractive index you calculated to the refractive index of the glass slab (assumed to be 1.5 unless otherwise stated).

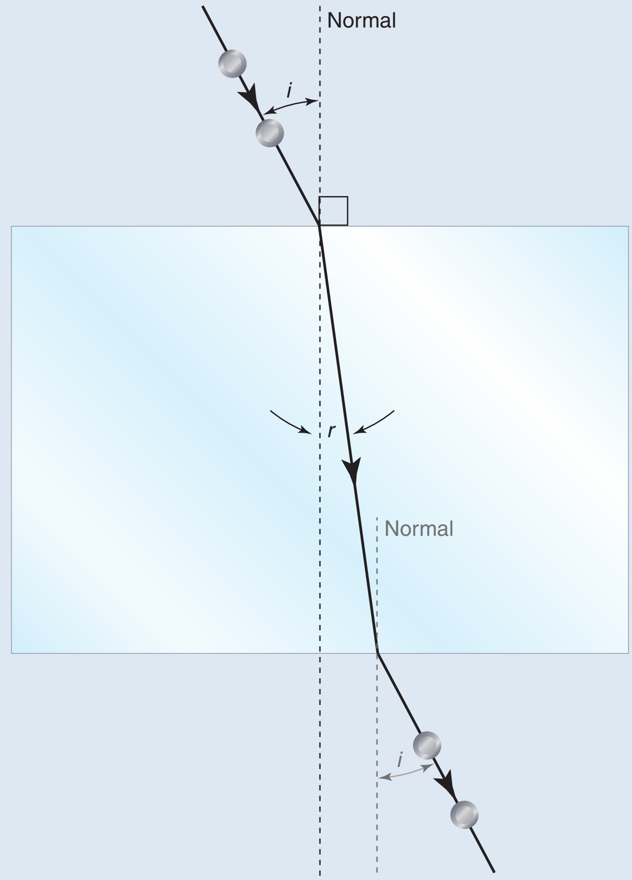


Figure 4.49 *The angles of incidence and refraction can be substituted into Snell's Law to determine the refractive index.*

Alternative method rather than using pins

Use an optics kit or a single, coherent light beam from a ray box or laser pointer and trace the path of the light beam into and out of the glass slab with a ruler. Proceed as above to calculate the refractive index of the slab.

Questions

1. Did your results compare with the published value of the refractive index of glass (1.5)? If they did not, propose reasons for any differences.
2. Have someone else perform the same experiment with the same slab of glass. Do your results compare?

PRELIMINARY MODULE

ELECTRICAL ENERGY IN THE HOME

CHAPTER 5

DISCOVERY AND DEVELOPMENT OF ELECTRICAL ENERGY

Figure 5.1 *Before the use of electric lighting, sporting events such as this could take place only in daylight hours.*

Remember

Before beginning this chapter, you should be able to:

- identify situations or phenomena in which different forms of energy are evident
- qualitatively account for the total energy involved in energy transfers and transformations, referring to the Law of Conservation of Energy.

Key content

At the end of this chapter you should be able to:

- identify the differing views of Galvani and Volta about animal and chemical electricity, and discuss whether their different views contributed to an increased understanding of electricity
- discuss how the main sources of domestic energy have changed over time
- assess some of the effects of changes in, and increased access to, sources of energy for a community
- discuss some of the ways in which electricity can be provided in remote locations.

In 1755, Samuel Johnson compiled his groundbreaking dictionary of the English language. The entry for *electricity* began:

> 'A property in some bodies, whereby, when rubbed so as to grow warm, they draw little bits of paper, or such like substances, to them . . . Such was the account given a few years ago of electricity, but the industry of the present age has discovered in electricity a multitude of philosophical wonders. Bodies electrified (may be given) such a quantity of electrical vapour, as, if discharged at once upon a human body, would endanger life. The philosophers are now endeavouring to intercept the strokes of lightning.'

Where Johnson's dictionary says *philosophical* we would say *scientific*. Where it says *electrical vapour* we would say *electric charge*. Making allowance for such slight differences in terminology, we can easily understand that the dictionary was referring to what is now called *static electricity*. It makes no mention of *electric currents*. Current electricity was not discovered until 1800. When Johnson compiled his dictionary, such a phenomenon was unknown.

In the two hundred years since their discovery, electric currents have become an all-pervading aspect of our way of life. In this chapter we will consider some of the discoveries that led to our present understanding of electric currents and look at how people have made use of electrical energy.

5.1 GALVANI AND VOLTA

In the eighteenth century, two men, Luigi Galvani and Alessandro Volta, investigated phenomena that led to the discovery of electric currents. Although Galvani and Volta had differing views about the explanation of these phenomena, they each contributed to an increased understanding of electricity.

Luigi Galvani

Luigi Galvani (1737–1798), was an Italian physician and physicist. His name is perpetuated in the name *galvanometer*, for a sensitive current detector, and in the term *galvanic cell*.

Figure 5.2 ***Luigi Galvani, Italian physician and physicist***

In 1772, Galvani acquired an electrostatic machine (a device for making sparks). Using dissected frogs' legs, he began to investigate the muscular contractions produced by electricity. Galvani found that a dissected frog's leg would contract if a scalpel touched the frog's nerves when the electrostatic machine was working or when there was a lightning strike in the vicinity.

In 1786, Galvani made a startling new observation. He found that he could produce contractions of a frog's leg even if there were no source of electricity present. If two different metals were attached to the dissected frog — one to a leg muscle and the other to the spinal cord — the leg muscle would contract if the two metals were touched together. (Galvani discovered this more or less accidentally when he suspended a dissected frog from an iron railing using a brass hook.)

Animal electricity was a term used by Galvani for a form of electricity that he believed was generated by animal tissues.

Galvani believed that his observations showed that electricity was being generated in the tissues of the frog. He called this type of electricity **animal electricity**. Some of Galvani's experiments are illustrated in figure 5.3.

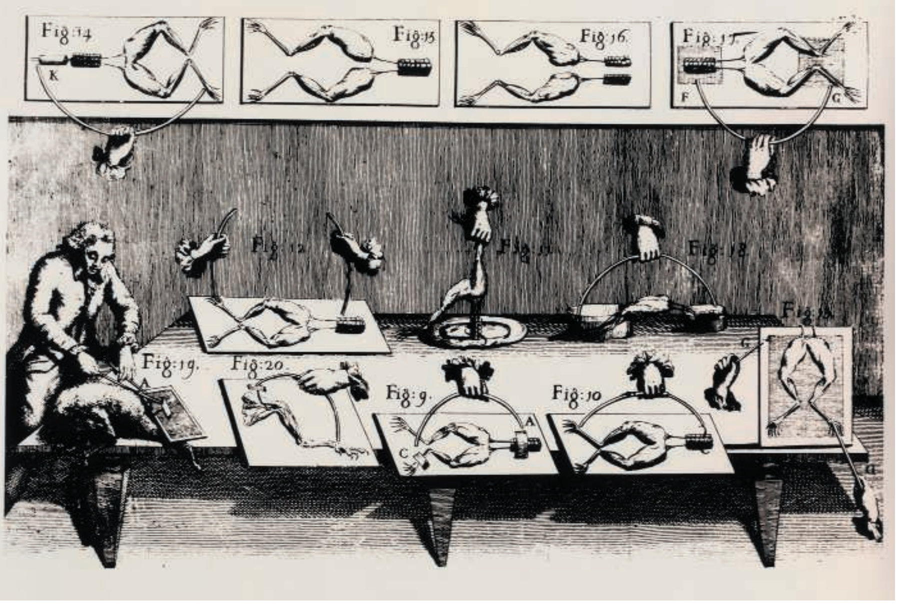


Figure 5.3 *Some of Galvani's experiments with dissected frogs' legs*

When Galvani published his theory of animal electricity in 1791, it caused a great stir in the scientific world. At the time it was commonly believed that there was an (as yet unknown) 'life force' that was responsible for endowing organisms with life. Speculation arose that Galvani's 'animal electricity' was the elusive 'life force'.

In a sense, Galvani was right. Animals do produce tiny electrical impulses that are essential for life, but these can only be detected by using sensitive amplifying equipment.

In modern terminology, Galvani thought that the muscle in the frog's leg produced a positive charge in one area of the muscle and a negative charge in another area. When the two areas were connected by a metal, a discharge took place which caused the muscle to contract.

Alessandro Volta

Alessandro Volta (1745–1827) was an Italian physicist. His name is perpetuated in the term *voltage* and its unit the *volt*.

Volta carried out experiments to confirm Galvani's work and at first accepted Galvani's theory of animal electricity. Galvani's theory, however, did not explain why two different metals were needed to produce contractions of the frog's leg. Volta became convinced that the source of the electricity was the contact between two different metals. He considered that the frog's leg contracted because the contact of two different metals produced electric charge. In Volta's view there was no animal electricity and no vital force. The frog's leg just acted as a detector of the electricity produced by the metals.

In modern terminology, Volta believed that when two different metals are brought into contact, one becomes positively charged and the other becomes negatively charged.

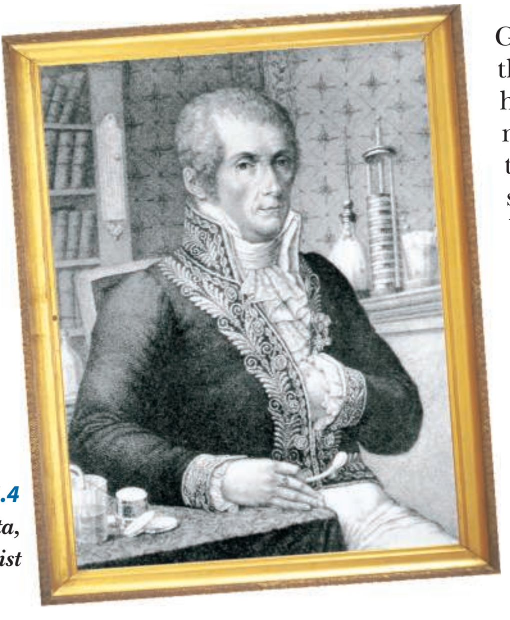


Figure 5.4 *Alessandro Volta, Italian physicist*

Volta tested his theory by using pairs of zinc and brass discs. To increase the effect he built up a pile, or battery, of as many as 60 of these pairs of discs. To improve the contact between the metal discs he put discs of cardboard moistened with salt solution between them. This assemblage of discs is referred to as a **voltaic pile** (see figure 5.5).

A voltaic pile is an assembly of large numbers of alternate zinc and brass discs separated by cardboard discs soaked in salt solution.

Volta joined conductors to the end discs of his pile. When he touched these conductors together sparks were produced. Moreover, the pile would continue to produce sparks each time the conductors were touched. As there was no animal involved this was taken as proof that Galvani's theory of animal electricity was wrong.

As well as cardboard, Volta also used other materials, such as felt and parchment, in making voltaic piles.

The voltaic pile was the first source of continuous electric current. It was the forerunner of present day batteries and its invention ushered in the modern electrical era.

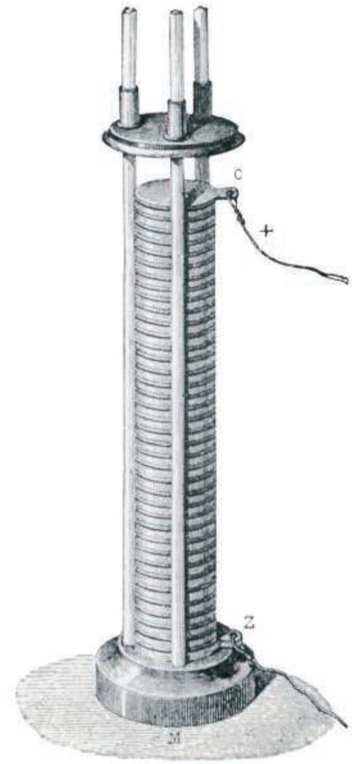


Figure 5.5 ***A voltaic pile***

Was Volta right?

Humphrey Davy (1778–1829) was an English chemist who discovered many elements including calcium and sodium. He invented the safety lamp which saved many lives by preventing explosions of flammable gases in mines.

Figure 5.6 ***Humphrey Davy, English chemist***

Davy showed that the source of the electricity produced by Volta's pile was actually chemical reactions involving the salt solution and the two different metals. Volta had believed that the salt solution was simply a way of making contact between the metals. Davy showed that the salt solution was essential to the production of the electricity. The electricity produced by the voltaic pile is **chemical electricity**.

Chemical electricity is the electricity produced by a chemical reaction.

However, Galvani had not used salt solution between the metals that he touched together to make the frog's leg contract. Davy's discovery that salt solution is essential in the voltaic pile raises the question: 'Why did the frog's leg contract when Galvani touched the metals together?'

The explanation is that when two different metals are brought into contact, one does become positively charged and the other does become negatively charged, exactly as Volta had thought. Volta believed he had proved this, but by placing salt solution between the metals he had actually discovered a completely different way of producing electricity.

Without the salt solution, contact between two different metals produces a pulse of current as the metals become charged. This pulse of current was enough to make the frog's leg contract, but was too weak to be detected in any other way by Galvani or Volta. With the salt solution, chemical reactions between the metals and the solution produce a continuous supply of electrical energy.

Volta set out to prove that when two different metals make contact, electricity is produced. He thought he had proved this, but in fact he proved that two different metals *with salt solution between them* produce electricity.

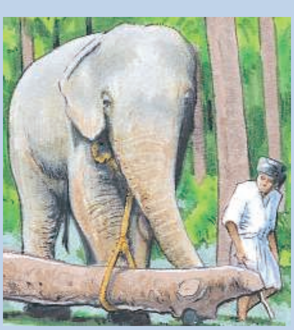
The production of electricity when two different metals make contact and the production of electricity when two different metals are separated by a salt solution are now regarded as different phenomena.

5.2 PEOPLE'S USE OF ENERGY SOURCES BEFORE ELECTRICAL ENERGY

People, like all animals, convert the chemical energy of food into heat energy in the body, and mechanical energy for movement and doing work. When this was not sufficient for the tasks they wanted to do, humans utilised other sources of energy. It is by using these other sources of energy that humans have had such an impact on the Earth.

Table 5.1 presents some information about energy sources used by people before electricity.

Table 5.1

ENERGY SOURCE	ABOUT THIS SOURCE OF ENERGY	SOME SOCIAL IMPACTS OF THIS SOURCE OF ENERGY
Wood 	• Chemical energy of wood converted into heat energy • First energy source used by humans • Used to produce fire	• Fire helped keep fierce animals at bay • Food could be cooked • Points of sticks could be hardened for weapons • Humans able to live in colder climates
Domesticated animals 	• Source of mechanical energy • Used to pull ploughs • Used to raise water from rivers and wells for irrigation • Used for transport • Used to turn grindstones • Used to lift heavy burdens	• More food could be produced, providing a surplus • Not everybody need be employed in producing food • Population increased • Food transported from farms to markets • Towns developed

ENERGY SOURCE	ABOUT THIS SOURCE OF ENERGY	SOME SOCIAL IMPACTS OF THIS SOURCE OF ENERGY
Wind and water 	• Source of mechanical energy • Wind energy used for boats • Windmills and water wheels used for grindstones	• Expanded trading • More efficient production of food from grain
Coal 	• Has been used as a fuel since about 1000 BC • Became main fuel in nineteenth century as a result of dwindling supplies of wood and superior energy content of coal	• Increased use of coal led to the development of the steam engine to pump water from coal mines • Higher energy content of coal enabled production of steel, an alloy of iron and carbon • Production of steel was one of the factors leading to modern industrial age • Manufacturing carried out in factories • Migration of workers from rural areas to cities • Subsequent overcrowding leading to development of slums • Spread of disease • Atmospheric pollution in cities
Coal gas 	• Fuel produced by heating coal in absence of air • Street lighting • Source of heat and light in houses	• Streets safer at night • Greater social activity at night

5.3 PEOPLE'S USE OF ELECTRICAL ENERGY

Generation of electrical energy

Although batteries based on the voltaic pile could produce large, continuous electric currents, these currents were mainly used for scientific research. The batteries were expensive and could only produce electric current for a limited time before the chemicals in the battery were exhausted. Before electricity could make any impact as a major source of energy, a cheaper and less limited method of generating electric currents had to be found.

Michael Faraday (1791–1867) was an English physicist and chemist. His name is perpetuated in the unit of capacitance, the *farad*. In 1831, Faraday discovered electromagnetic induction. This led to the development of the **electric generator** in which electrical energy is produced by rotating a coil of wire in a magnetic field. This is the basis of the large-scale production of electrical energy in power stations.

An electric generator is a device in which electrical energy is produced by rotating a coil in a magnetic field.

Electromagnetic induction is studied in the HSC course.

In order to produce electrical energy, some other source of energy is required to rotate the coil in the generator.

The most common power stations are **fuel-burning power stations**. They produce heat energy by burning fossil fuels, most commonly coal, but also oil and natural gas. The heat energy is used to boil water, producing high-pressure steam which passes through a turbine and causes it to rotate. The coil of the generator is connected to the shaft of the turbine and rotates with it. Figure 5.8 shows a coal-burning power station.

A fuel-burning power station is one in which fossil fuels are burnt to provide energy.

Figure 5.7 ***Michael Faraday, English physicist and chemist***

Figure 5.8 ***A coal-burning power station***

Figure 5.9 *A nuclear power station*

A nuclear power station is one in which nuclear reactions provide energy.

A hydro-electric power station is one in which water that has gained kinetic energy by flowing downhill is used to provide energy.

Other forms of power station include:

- **nuclear power stations**, where nuclear energy is used to produce heat energy. This heat energy is used in exactly the same way as in a fuel-burning power station. Figure 5.9 shows a nuclear power station.
- **hydro-electric power stations**, where the energy required to turn the turbine is provided by a stream of water which has gained kinetic energy by flowing downhill, usually from a dam (see figure 5.10). George Westinghouse built the first hydro-electric power station using water in Niagara Falls.

Figure 5.10 *A hydro-electric power station*

Distribution of electrical energy

One of the great advantages of electrical energy is that it can be transferred almost instantaneously from the power station where it is produced to the communities where it is used. Transmission lines are used to carry the electrical energy across great distances.

In a typical power station the voltage generated is around 10 000 volts and the current produced is around 10 000 amps. A current of this magnitude would waste so much electrical energy as heat in the transmission wires that the use of electrical energy would be uneconomical. Also, a voltage of 10 000 volts would be dangerous to use.

Transformers are studied in the HSC course.

These difficulties can be overcome by the use of transformers. These are devices that can change the voltage and current to suitable values in different parts of the distribution system.

Figure 5.11 shows a typical distribution system with the voltages in various parts of the system. Note the very high voltage in the transmission lines. This high voltage means that the current through the transmission lines is very low and therefore, that the energy wasted in the transmission lines will be very small.

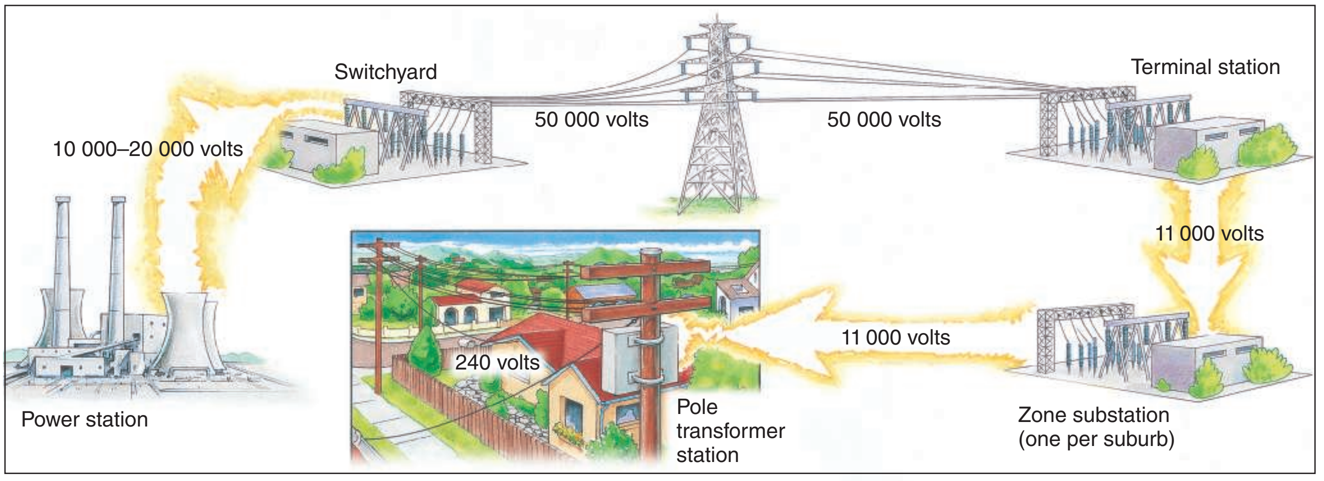
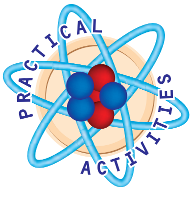


Figure 5.11 *A typical distribution system*

Spread of use of electrical energy

Although Faraday discovered the fundamental principles behind the generation of electricity in 1831, it was not until the invention of the incandescent lamp by Thomas Edison in 1879 that a real stimulus was given to the use of electrical energy.

To highlight the rapid spread of electrical energy after the discovery of the incandescent lamp, consider the following dates illustrating the introduction of electrical energy in New South Wales:



5.1

Use of electrical energy in Australia

- 1888 Tamworth the first town in New South Wales to have electric street lights; Young followed later in the year
- 1889 Street lighting in Penrith and Moss Vale
- 1890 Street lighting in Broken Hill
- 1891 Street lighting in Redfern
- 1892 Street lighting in Newcastle
- 1902 First electric street lights in the City of Sydney installed in King Street between Elizabeth and Pitt streets
- 1904 First public power station built at Pyrmont with a total capacity of 1500 kilowatts

- 1905 There were 519 customers connected. (Many customers were unable to use electrical energy in their factories because of the cost of the motors. A scheme to hire motors was introduced and two years after its commencement 361 motors were on hire.)
- 1910 Decision made to change completely to electric public lighting in the City of Sydney
- 1910 Expansion into the suburbs on the southern sides of Sydney Harbour started
- 1913 Almost 10 000 customers, with total consumption of 40 000 000 units of electricity
- 1914 Two submarine cables laid across the bed of Sydney Harbour to supply the northern side of the harbour.

The electricity grid

As the use of electrical energy and the number of power stations increased, a group of power stations were linked together in a grid. Power stations connected in this way can exchange energy so that a station with a low demand for energy at a particular time can assist one experiencing a high demand. Most of eastern Australia will eventually be connected in a single grid. Remote towns, such as Mt Isa, are not connected to the main grid but have their own power stations.

It is possible to feed electrical energy into the electricity grid from sources that would not, by themselves, provide sufficient electrical energy. For example, banks of many windmills, called wind farms, have been constructed in places where there are steady high winds. The energy of the wind is used to rotate a turbine connected to the coil of a generator. Figure 5.12 shows a wind farm.

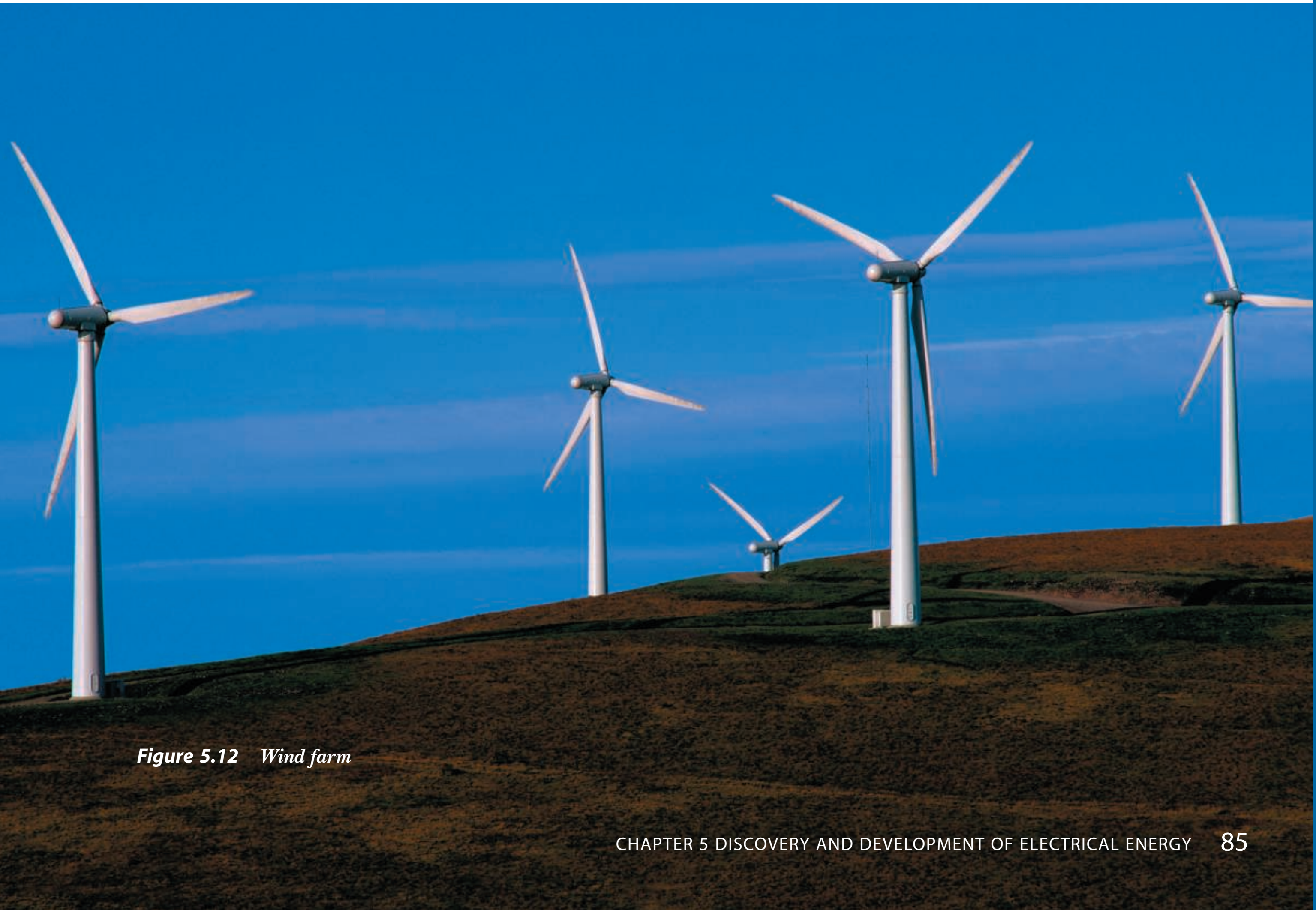


Figure 5.12 *Wind farm*

Electricity in remote places

There are many places too remote to be connected to a power station. Such places frequently use small generators in which the coil is rotated by an internal combustion engine using petrol or oil. Arrays of solar cells, which convert the energy of sunlight into electrical energy, and small wind generators are also used for some specialised purposes (see figures 5.13 and 5.14).

Figure 5.13 *A small wind generator*

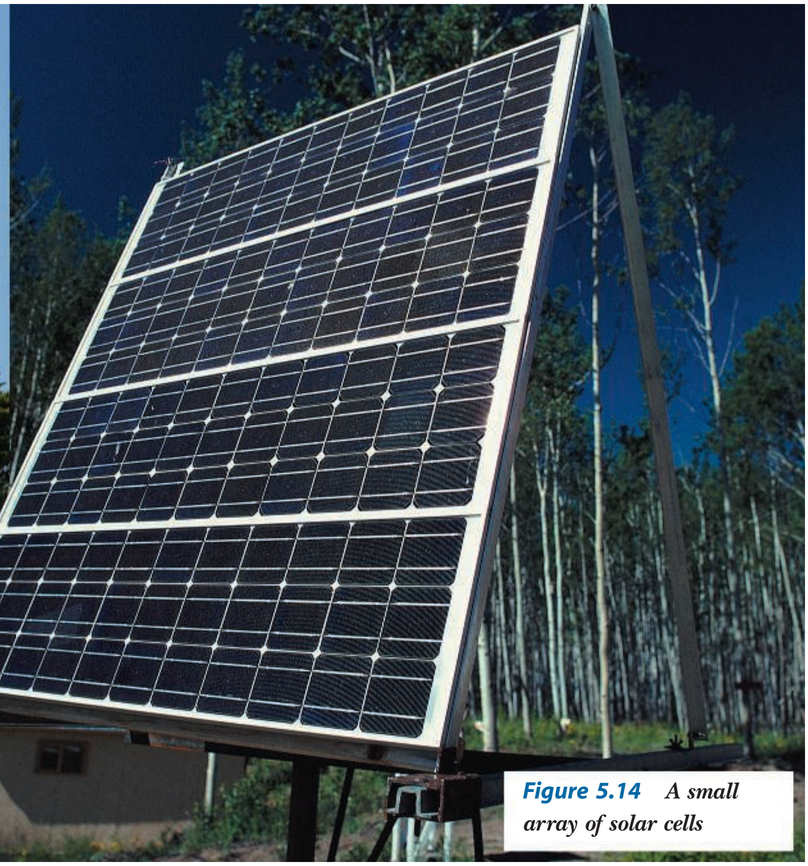


Figure 5.14 *A small array of solar cells*

Social impact of the use of electrical energy

As it is now possible to use electrical energy to provide light as bright as daylight, activities that were once confined to the daylight hours can now be carried on at any time. Many factories run continuously for 24 hours a day. Some shops are open until late or all night. Sporting events can take place in the evening when many more people can attend.

Before the advent of electric heaters, cities such as London were troubled with smog from the burning of wood and coal. Although coal is still burnt in the power stations, it is possible to burn the coal more completely and minimise the amount of polluting material that reaches the atmosphere.

The development of small electric motors in the 1920s led to the introduction of many domestic appliances. Many of these are now regarded as standard in most households. Refrigerators have enabled food to be stored for long periods and have lessened the need for daily shopping. Washing machines and clothes dryers have reduced the time needed for cleaning clothes. Reductions such as these in the time spent on housework have been among the factors in the movement of women into the workforce.

PHYSICS IN FOCUS

The Snowy Mountains Hydro-electric Scheme

Every winter, snow accumulates on the Australian Alps. When the snow melts each spring, melt-water flows eastward to the sea through the Snowy River and its tributaries. The Snowy Mountains Hydro-electric Scheme (usually referred to as the Snowy Mountains Scheme) diverts much of this water westward for irrigation.

At the heart of the Scheme is Lake Eucumbene, formed by damming the Eucumbene River, a tributary of the Snowy River. Lake Eucumbene is the Scheme's largest reservoir, with a volume about nine times that of Sydney Harbour. Another large reservoir is Lake Jindabyne, formed by damming the Snowy River.

From these reservoirs, water passes through huge underground tunnels that take it westward under the mountains. The Scheme has two branches, a north-west branch leading to the Murrumbidgee River and a south-west branch leading to the Murray River. A map of the pipelines is shown in figure 5.16 on page 88. (This map omits many of the tributaries of the rivers involved in the Scheme.)

In its passage through the system the water falls through a height of 800 metres and, as it falls, it passes through a number of hydro-electric power stations generating electrical energy. The Scheme generates nearly 3880 MW of electrical power. Figure 5.17, on page 88, shows one of the power stations of the Snowy Mountains Scheme.

Work on the Scheme began in 1949 and took nearly 25 years to complete. The Scheme comprises 16 large dams and many smaller ones, 145 km of mountain tunnels, 7 power stations (2 of them underground), over 80 km of high mountain aqueducts, 1600 km of roads and hundreds of kilometres of transmission lines.

(continued)

Figure 5.15 ***Lake Eucumbene and Eucumbene Dam***

Although the Scheme was very successful in generating electrical energy and in providing water for irrigation west of the dividing range, it has become clear in recent years that there has been an environmental cost. For example, the reduced water flow through the Snowy River has contributed to problems of salination that threaten the agricultural areas through which the river passes. Measures are being taken to increase the flow of water through the Snowy River to try to overcome this problem.

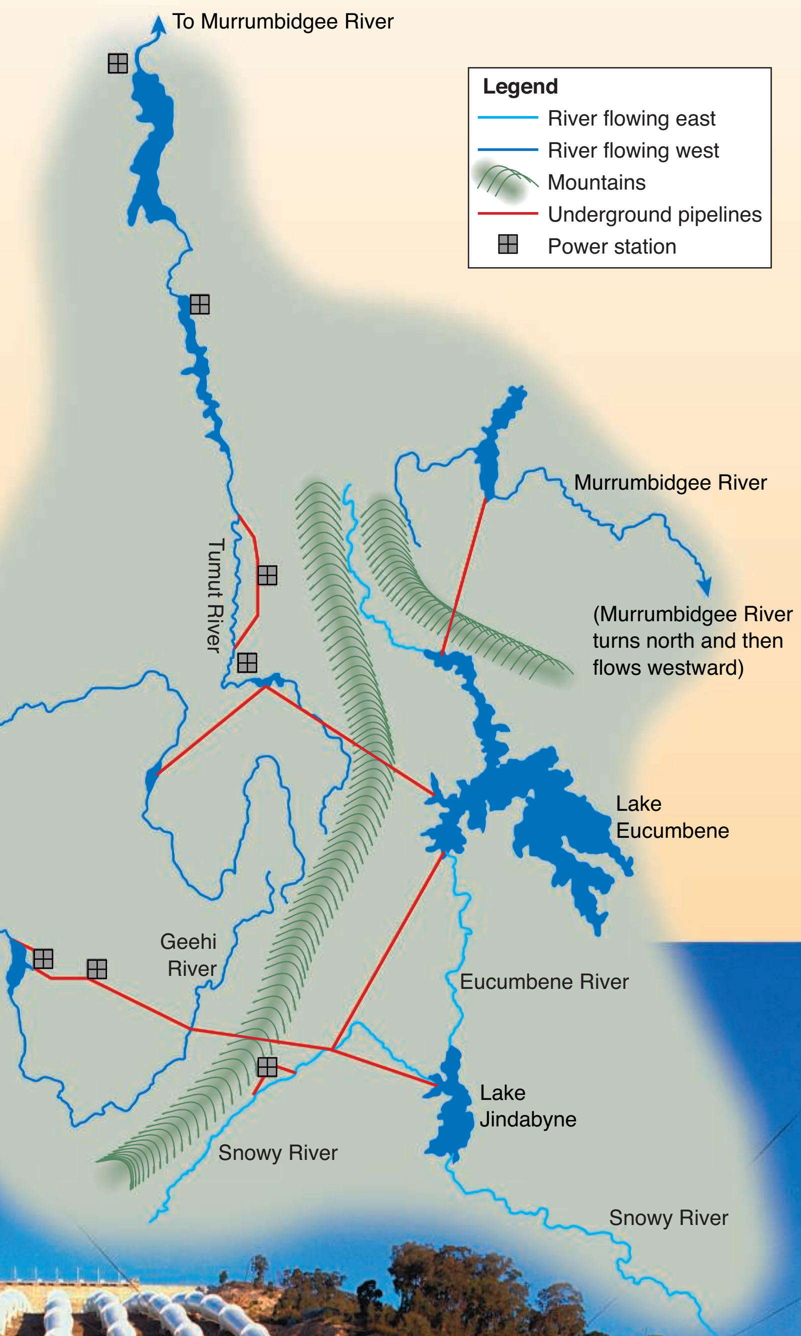


Figure 5.16 *A simplified map showing the underground pipelines of the Snowy Mountains Scheme*

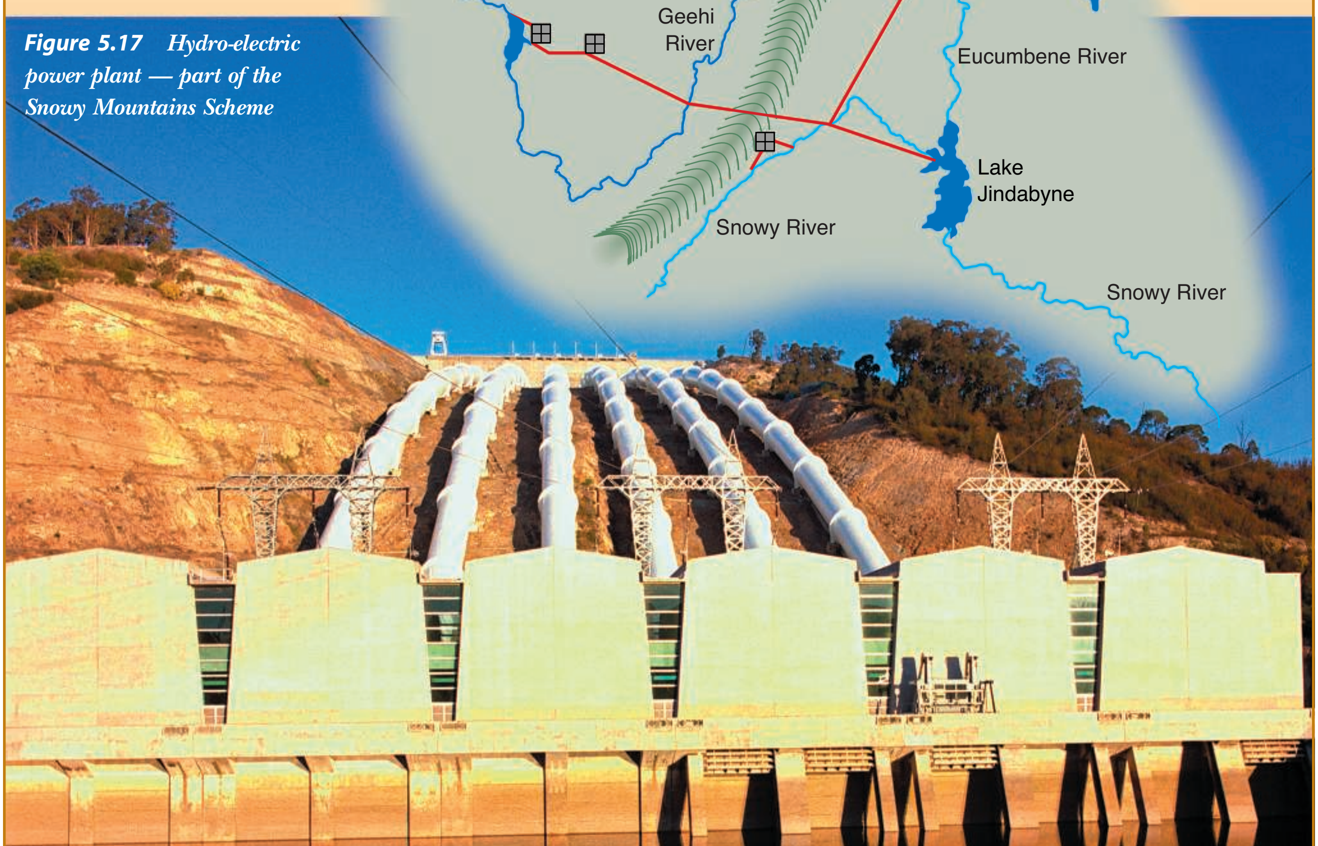


Figure 5.17 *Hydro-electric power plant — part of the Snowy Mountains Scheme*

SUMMARY

- Galvani observed that two different metals attached to a dissected frog caused the frog's leg to contract when the two metals were touched together.
- Galvani believed that the electricity that caused the frog's leg to contract was produced in the frog's leg. He called it 'animal electricity'.
- Volta believed that the electricity was produced by the contact of the two different metals. He produced an electric current using two different metals separated by salt solution. This is called 'chemical electricity'.
- Volta invented the 'voltaic pile', made of alternate discs of zinc and brass separated by discs of a material, such as cardboard, soaked in salt solution. This produced the first continuous source of electric current. The voltaic pile is the forerunner of present-day batteries.
- The different views of Galvani and Volta contributed to an increased understanding of electricity.
- Humans have utilised many sources of energy for their needs through the ages, enabling them to increase in numbers and form complex, mainly urban societies.
- After the invention of the incandescent light bulb, the use of electrical energy spread quickly. There has been a continuous rapid rise in the amount of electrical energy used.
- The fuel-burning power station is the most common type for the large-scale generation of electrical energy. Other types are nuclear power stations and hydro-electric power stations.
- By means of transmission lines, the electrical energy produced in a power station can be distributed almost instantaneously to the consumer.
- A grid of transmission lines can be built up with many power stations feeding electrical energy into the grid. Smaller electrical energy generators can also feed energy into the grid.
- Small generators have been developed to provide electrical energy to remote areas. Small wind generators and arrays of photoelectric cells are also used.
- The use of electrical energy has had a profound and often revolutionary impact on people, and their communities, in many areas of their lives.

QUESTIONS

1. (a) Explain the meaning of the term 'animal electricity'.

   (b) Describe the observation that led Galvani to postulate the existence of animal electricity.
2. (a) Explain the meaning of the term 'chemical electricity'.

   (b) Describe the observation that led Volta to reject Galvani's theory of animal electricity.
3. Discuss whether the different views of Galvani and Volta led to an increased understanding of electricity.
4. Outline the structure of Volta's 'pile'. Discuss its significance for the development of current electricity.
5. Identify three of the sources of energy used by human beings before 1850. For each source assess the impact on the community.
6. Discuss the significance of Edison's invention of the incandescent light bulb for the spread of electricity as an energy source.
7. Discuss, in an essay or a talk, the impact of the use of electrical energy on individuals and on the community.
8. Make a list of the number of ways you used electrical energy yesterday.
9. Make a list of the appliances in your home that use electric motors. For each appliance discuss:

   (a) what work is done by the appliance

   (b) how this work would have been done before the use of electricity.
10. Discuss how you would be affected if there were a shutdown of the electricity supply for (i) an hour, (ii) a day, and (iii) a week.
11. Discuss the increase in the use of electrical energy since it first became available.
12. Discuss the use of wind farms in the production of electrical energy.
13. State three ways in which electricity can be provided to remote locations.

5.1 USE OF ELECTRICAL ENERGY IN AUSTRALIA

Aim

To investigate the growth in the use of electrical energy in Australia in the period 1920–1980

Method

This is an activity involving interpretation of data. The first table shows electrical energy consumption in Australia between 1920 and 1980. The second table below shows the population of Australia over the same period.

Using the information in the tables draw the following graphs:

1. electrical energy consumption against time
2. population against time
3. electrical energy consumption per head of population against time.

Analysis

1. What was the percentage increase in the population from 1920 to 1980?
2. What was the percentage increase in electricity consumption during this time?
3. Comment on these results.
4. From these figures, estimate the population and electrical energy consumption in 2000.

Questions

1. Is it valid to estimate as you did?
2. Give reasons for your view.
3. Find out the population of Australia in 2000 and compare it with the result of your estimate.

Electricity consumption for Australia

YEAR	CONSUMPTION (GWh)
1920	487
1930	2 436
1940	5 180
1950	8 475
1960	21 449
1970	49 412
1980	87 328

Population of Australia

YEAR	POPULATION
1921	5 435 734
1933	6 629 839
1947	7 579 358
1954	8 986 530
1961	10 508 186
1966	11 550 462
1971	12 755 638
1976	13 548 450
1981	14 576 330

CHAPTER 6

ELECTRIC CHARGES, FIELDS AND CURRENTS

Figure 6.1 Lightning is a naturally occurring example of electrical phenomena.

Remember

Before beginning this chapter, you should be able to:

- recall that all matter consists of atoms
- state that most of the mass of an atom is in the nucleus
- describe ways in which objects acquire an electrostatic charge
- describe the behaviour of charges when brought close to each other
- describe voltage, resistance and current using analogies
- describe qualitatively the relationship between voltage, resistance and current.

Key content

At the end of this chapter you should be able to:

- describe the behaviour of electrostatic charges and the properties of the fields associated with them
- define the unit of electric charge as the coulomb
- define the electric field as a force field with a field strength at a point equal to the force per unit charge placed at the point
- define the direction of the electric field at a point as the direction of the force on a small positive charge placed at the point
- solve problems and analyse information using $\boldsymbol{E} = \frac{\boldsymbol{F}}{q}$
- present diagrammatic information to describe the electric field about and between a positive and negative point charge and between oppositely charged parallel plates
- define electric current as the rate at which charge flows under the influence of an electric field
- define the unit of electric current as the ampere and identify that one ampere is equivalent to one coulomb per second
- identify that current can be either direct or alternating
- describe electric potential difference (voltage) between two points as the change in potential energy per unit charge moving from one point to the other
- identify the unit of electric potential difference (voltage) as the volt and that one volt is equivalent to one joule per coulomb
- discuss how potential difference changes at different points around a DC circuit
- identify the difference between conductors and insulators
- define resistance as the ratio of voltage to current for a particular conductor, and solve problems using $R = \frac{V}{I}$
- describe qualitatively how the movement of electricity through a conductor is affected by length, cross-sectional area, temperature and material
- identify conductors commonly used to provide household electricity
- identify uses of voltmeters and ammeters and explain why they are connected differently in a circuit.

6.1 ELECTRIC CHARGE

The words *electric* and *electricity* are derived from the Greek word for amber: *electron*. Amber is a naturally occurring substance exuded as a resin from certain trees. As long ago as 500 BC, the Greeks had observed that if amber was rubbed it would attract small pieces of material. Today we can observe this phenomenon more conveniently using certain man-made materials such as perspex. When a perspex rod is rubbed with silk, the rod acquires the ability to attract small pieces of materials such as paper. The rod is said to have become electrically charged.

Some other common observations of bodies becoming electrically charged are:

- when you walk on a carpet on a dry day your body becomes electrically charged. If you touch a metal door handle you feel a slight shock as your body is discharged.
- on a dry day a car becomes electrically charged as it moves through the air. If you touch the car you feel a slight shock as the car discharges through your body.

Electric charge and the structure of atoms

We now understand electric charge in terms of the basic structure of matter. All matter is made of atoms that are themselves made of electrons, protons and neutrons, as shown in figure 6.2.

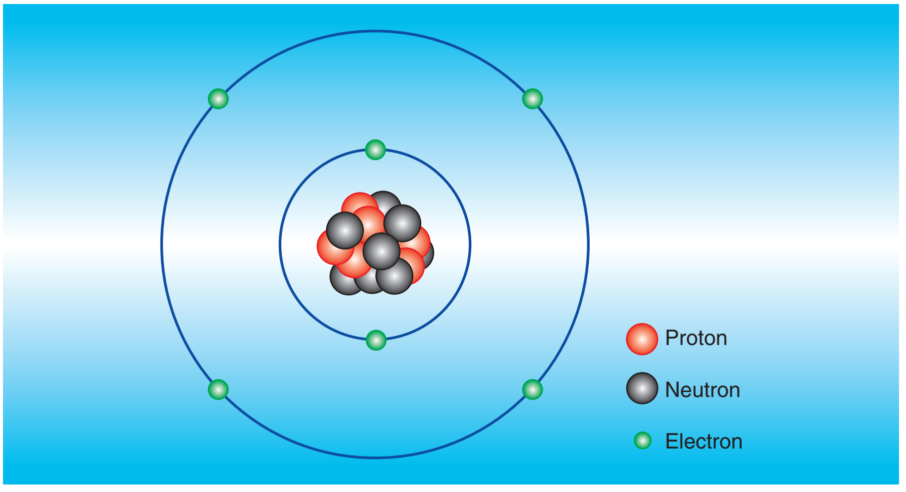


Figure 6.2 *The structure of an atom*

Electric charge is a property of electrons and protons by which they exert electric forces on one another.

Positive charge is the type of charge on a proton.

Negative charge is the type of charge on an electron.

Electric charge is a property of electrons and protons. Because of their electric charge, these particles exert electric forces on each other. Protons carry a **positive charge**; electrons carry a **negative charge**. The positive charge on a proton is equal in magnitude to the negative charge on an electron.

The directions of the forces between electric charges are:

- two positive charges repel one another
- two negative charges repel one another
- a positive charge and a negative charge attract one another.

This is summarised as: *like charges repel; unlike charges attract.*

Neutrons, the third type of particle in atoms, have no electric charge and do not experience electric forces. Neutrons are uncharged or neutral.

The **coulomb** (C) is the SI unit of electric charge.

The name, **coulomb**, *for the SI unit of electric charge is derived from Charles Augustin Coulomb (1736–1806), a French physicist who studied the forces between electric charges.*

The SI unit of electric charge is the **coulomb** (C). The coulomb is defined in terms of electric current, but for our purposes, it is sufficient to state that a charge of one coulomb is approximately equal to the total charge on 6.25×10^{18} electrons (or 6.25×10^{18} protons). That is, the charge on one electron is approximately -1.60×10^{-19} C, and the charge on one proton is approximately $+1.60 \times 10^{-19}$ C.

A coulomb is a very large charge; two charges of 1 C placed 1 metre apart would exert forces on each other of approximately 10^{10} N. A smaller unit of charge, the microcoulomb (μC), is often used. ($1\ \mu C = 10^{-6}$ C)

The symbols Q and q are usually used to represent electric charge. For example: $Q = 1.4 \times 10^{-5}$ C.

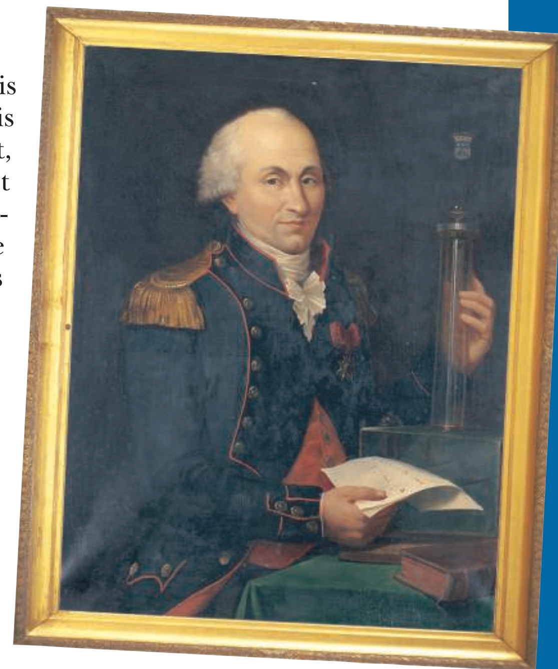


Figure 6.3 *Charles Augustin Coulomb, French physicist*

Neutral and charged bodies

If a body is **neutral** it has equal numbers of protons and electrons.

An **excess of electrons** exists when a body has more electrons than protons.

A **negatively charged** body has an excess of electrons.

A **deficiency of electrons** exists when a body has fewer electrons than protons.

A **positively charged** body has a deficiency of electrons.

An **electrostatic charge** is a charge due to an excess or deficiency of electrons.

A body that has equal numbers of protons and electrons will be **neutral**. As each atom in a body normally has equal numbers of protons and electrons, most bodies are neutral. However, it is possible for a body to lose some of its electrons or to gain extra electrons.

If a body has *gained* electrons it will have more electrons than protons. The body has an **excess of electrons** and is **negatively charged**. If a body has *lost* electrons it will have fewer electrons than protons. The body has a **deficiency of electrons** and is **positively charged**. A charge on a body due to an excess or deficiency of electrons is called an **electrostatic charge**.

It is always electrons that are gained or lost by a body, as the protons are strongly bound in the nuclei at the centres of the atoms. For this reason we talk about excess and deficiency of electrons rather than deficiency and excess of protons.

The deficiency or excess of electrons in a charged body is only a minute fraction (typically no more than 1 in 10^{12}) of the total number of electrons in the body. When we refer to the charge on a body, it is always the net charge that is meant.

SAMPLE PROBLEM 6.1

Calculating number of electrons for a given charge

A body has a charge of $+4.60\ \mu$C.

(a) Does it have an excess or a deficiency of electrons?
(b) Calculate how many excess or deficient electrons the body has.

SOLUTION

(a) Since the body is positively charged, it has a deficiency of electrons.
(b) As the charge on one electron is 1.60×10^{-19} C, the number of deficient electrons, n, that have a charge equal to $4.60\ \mu$C is given by:

$$n = \frac{(4.60 \times 10^{-6})}{(1.60 \times 10^{-19})}$$
$$= 2.88 \times 10^{13}$$

Conductors and insulators

A **conductor** is a material that contains charge carriers.

A **charge carrier** is a charged particle that is free to move through a material.

An **insulator** is a material that does not contain charge carriers.

A body is **insulated** if it is not earthed.

A body is **earthed** if it is connected to the Earth by a conducting path.

A **conductor** is a material that contains **charge carriers**; that is, charged particles which are free to move through the material. Examples of conductors and their charge carriers are:

- salt solutions — the charge carriers are positive and negative ions that are free to move through the solution
- metals — the charge carriers are electrons.

In this preliminary course, metals will be the only conductors studied.

An **insulator** is a material that contains no charge carriers. Common insulating materials are dry air, glass, plastics, rubber and ceramics. If an insulator is given an electrostatic charge at a particular area on the insulator, the charge will remain at that area.

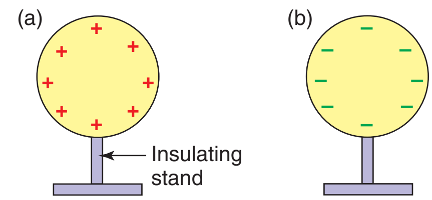


Figure 6.4 *Charge on insulated conductors (a) Charge on a positively charged conductor and (b) charge on a negatively charged conductor*

If a conductor is given an electrostatic charge, there are two possibilities:

- If the conductor is **insulated** (not earthed), there will be a movement of electrons within the conductor so that the electrostatic charge is as widely spread as possible. The electrostatic charge will be distributed on the surface of the conductor (see figure 6.4).
- If the conductor is **earthed**; that is, if there is a conducting path between the conductor and the Earth, electrons will move to or from the Earth to neutralise the conductor. See figure 6.5 (note the symbol for an earth connection). The Earth is so big that the negative charges going to or leaving the Earth produce no detectable charge on it.

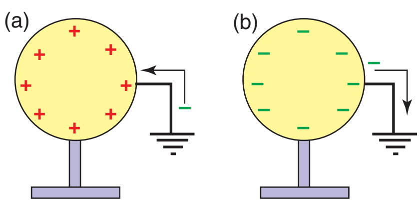


Figure 6.5 *(a) Earthing a positively charged conductor and (b) earthing a negatively charged conductor*

Charging by friction

The term 'charging by friction' is not really accurate. It is the contact *between two different materials that brings about the transfer of electrons; rubbing a solid body with a cloth is simply a way of improving the contact between them.*

If two bodies made of different materials are rubbed together, a small number of electrons will be transferred from one body to the other. The body that has lost electrons will have a deficiency of electrons and will be positively charged. The body that has gained electrons will have an excess of electrons and will be negatively charged.

The direction in which electrons are transferred depends on what two materials are rubbed together. It is possible to list materials so that when two of them are rubbed together, the top-listed material becomes positively charged and the bottom-listed material becomes negatively charged. A partial list of this type is given at left.

Air
Rabbit fur
Glass
Human hair
Nylon
Wool
Silk
Steel
Wood
Perspex

For example:

- If glass is rubbed with silk, electrons are transferred from the glass to the silk (see figure 6.6).

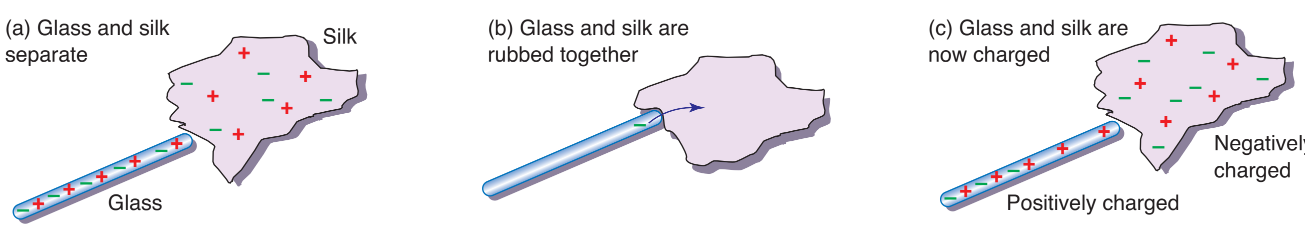


Figure 6.6 *Positively charging glass by friction*

To charge a metal by friction it must be insulated. Otherwise it will immediately lose any charge it acquires.

- If perspex is rubbed with wool, electrons are transferred from the wool to the perspex (see figure 6.7).

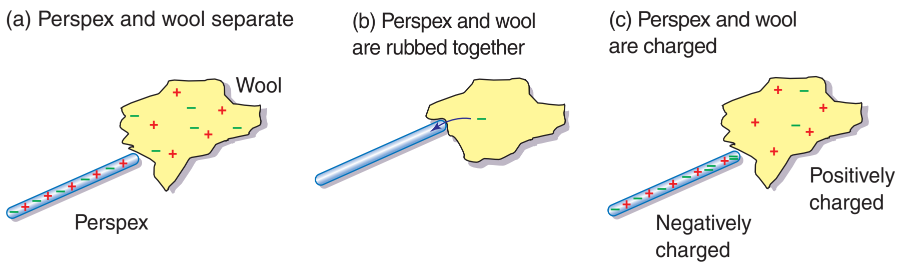


Figure 6.7 *Negatively charging perspex by friction*

Charging by contact

If a charged conductor is brought into contact with an uncharged conductor, the charge will be shared between the two conductors. The uncharged conductor will be charged by contact.

Figure 6.8 shows a neutral conductor being charged by contact with a positively charged conductor. Figure 6.9 shows a neutral conductor being charged by contact with a negatively charged conductor.

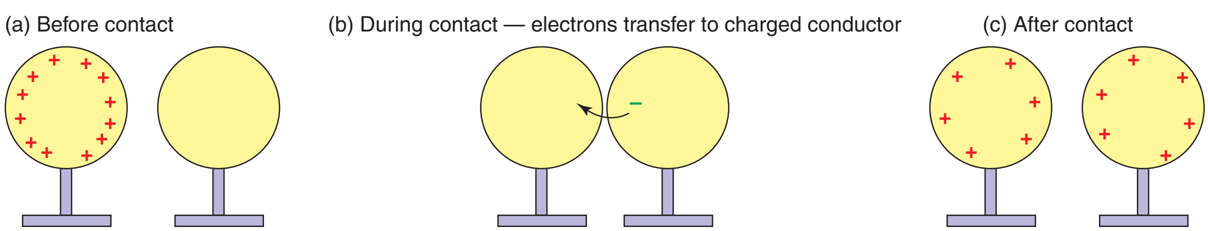


Figure 6.8 *Charging by contact with a positively charged body*

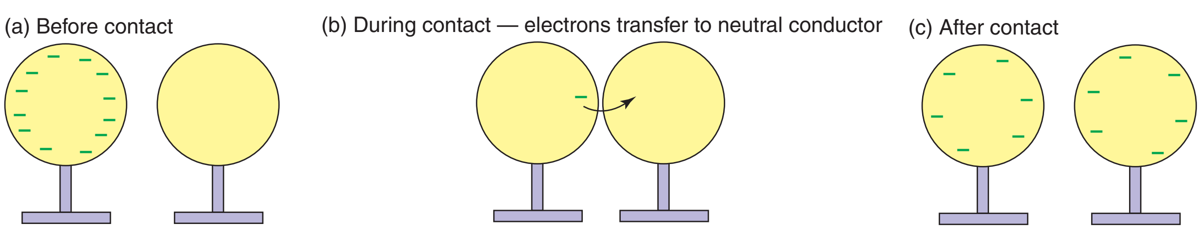


Figure 6.9 *Charging by contact with a negatively charged body*

Induced charges

In figures 6.8 and 6.9 the metal spheres are on insulating stands.

As an example of induced charges, consider what happens when a positively charged body is brought near an insulated, uncharged conductor. The positively charged body will attract electrons in the conductor. Some of these electrons will move to the area of the conductor closest to the positively charged body. As a result, that end of the conductor will have an excess of electrons (be negatively charged) and the opposite end of the conductor will have a deficiency of electrons (be positively charged). The charges on the conductor are called **induced charges** and the process is called **induction**. This is illustrated in figure 6.10.

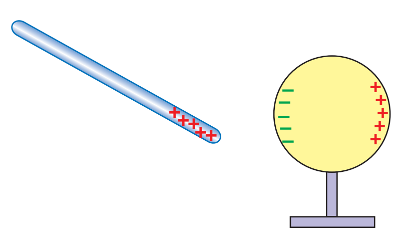


Figure 6.10 *Induced charges in an insulated conductor*

Induced charges are charges produced in a body when another charged body is near it.

Induction is the production of induced charges.

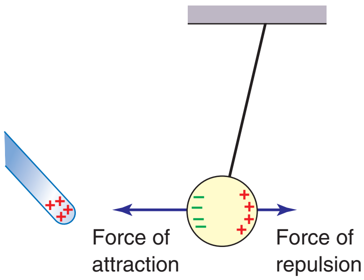


Figure 6.11 ***Attraction between a charged body and a neutral conductor***

Because they are closer, the positively charged body attracts the negative induced charges more strongly than it repels the positive induced charges. There will be a net force of attraction between the positively charged body and the conductor. This is illustrated in figure 6.11.

Induced charges are also produced if a charged body is brought near an insulator. For example, a positively charged body will attract the electrons and repel the nuclei in each atom of the insulator. These forces of attraction and repulsion result in a slight separation of positive and negative charges within each atom. As a result, one end of the insulator will be negatively charged and the other end will be positively charged, as illustrated in figure 6.12.

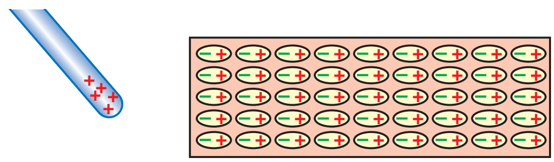


Figure 6.12 ***Induced charges in an insulator***

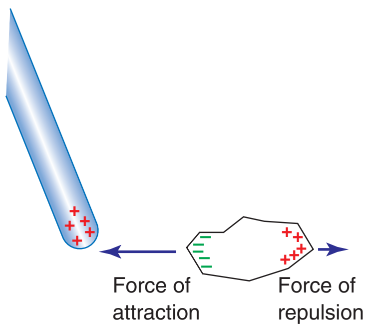
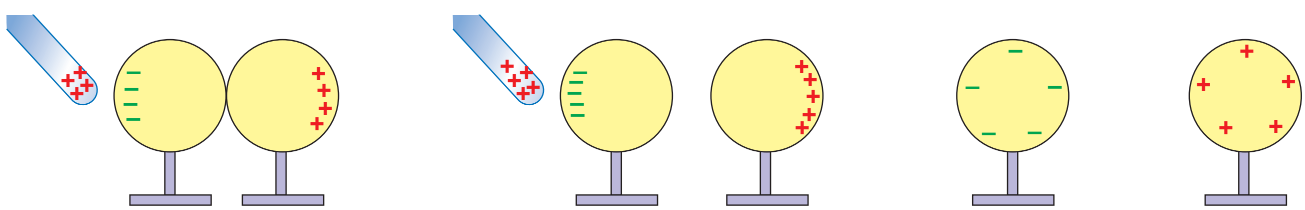


Figure 6.13 ***Attraction between a charged body and a neutral insulator***

Induction explains why a charged body attracts small uncharged bodies such as pieces of paper. This is illustrated in figure 6.13.

Permanently charging by induction

Induced charges are usually not permanent; when the charging body is removed the induced charges disappear. However, it is possible to charge an insulated conductor permanently by induction. One method of doing this is shown in figure 6.14.



(a) Bring the charged body near two touching, insulated conductors.

(b) Move the insulated conductors apart.

(c) Remove the charged body.

Figure 6.14 ***Charging a conductor permanently by induction***

Conservation of charge

When two previously neutral bodies are charged by friction, the amount of positive charge produced on one body is equal to the amount of negative charge produced on the other body.

When two charged conductors are brought into contact, there is a redistribution of charge between the bodies but the total amount of charge remains the same.

Observations such as these lead to the conclusion that the total amount of electric charge never changes; that is, electric charge is conserved.

SAMPLE PROBLEM 6.2

Conservation of charge

Two identical, insulated metal spheres carry charges of +3.0 μC and −7.0 μC. The spheres are brought into contact and then separated. Calculate the new charge on each sphere.

SOLUTION The total charge is −4.0 μC. As the spheres are identical, this charge will be shared equally by the two spheres. Therefore, the charge on each sphere will be −2.0 μC.

PHYSICS IN FOCUS

Lightning

Lightning is a natural phenomenon that illustrates some of the properties of electric charge.

Processes inside a storm cloud cause the bottom of the cloud to become negatively charged and the top to become positively charged. The mechanism responsible for this separation of charge is not known for certain, but many scientists think that the following happens. The inside of the cloud contains minute particles of ice. When two ice particles of different sizes collide, there is a transfer of electrons, so that the smaller particle becomes positively charged and the larger particle becomes negatively charged. Under the influence of gravity and updraughts within the cloud, the larger, negatively charged ice particles move towards the bottom of the cloud and the smaller, positively charged ice particles move towards the top of the cloud. Charges in the order of a coulomb can accumulate in this way.

The negative charge at the bottom of the cloud repels electrons from the Earth's surface. The ground under the cloud therefore becomes positively charged by induction.

Air is normally an insulator. Before lightning can occur, the air must become conducting. The negative charges at the bottom of the cloud and the positive charges on the ground under the cloud exert strong forces on the electrons and nuclei of atoms of air between the cloud and the Earth. The electrons experience forces towards the Earth; the nuclei experience forces towards the cloud. These forces can be so great that some electrons are removed from their atoms. An atom that has lost one or more electrons is positively charged. The liberated electrons and the positively charged atoms act as charge carriers, so the air becomes conducting. Once this process begins, a complicated chain of events follows which leads to the establishment of a conducting path between the cloud and the Earth. Immediately after the conducting path is complete there is a flow of negative charge from the cloud to the ground. This is a lightning flash and is shown in figure 6.1. A lightning flash usually consists of about four separate strokes, each stroke lasting about 30 microseconds.

The air along the path of a lightning stroke is heated to a temperature of about 20 000°C, producing light. The heated air expands producing a shock wave in the surrounding air that is heard as thunder.

As well as occurring between a cloud and the Earth, lightning also occurs within a cloud and between two clouds. Figure 6.15 illustrates different paths that lightning may travel.

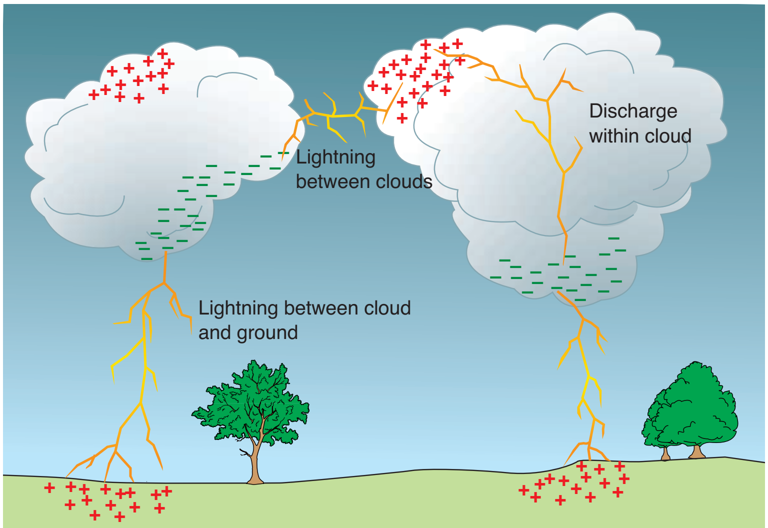


Figure 6.15 *Different paths for lightning*

6.2 ELECTRIC FIELDS

Field model of electric forces

Up to now we have considered a model where two electric charges exert forces directly on one another as shown in figure 6.16.

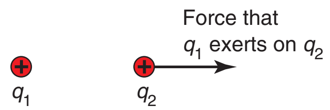


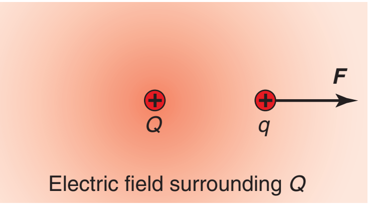
Figure 6.16 *An electric charge exerting a force directly on another charge*

An **electric field** is a field of force with a field strength equal to the force per unit charge at that point: $\boldsymbol{E} = \frac{\boldsymbol{F}}{q}$.

The term 'field' refers to a quantity that has a value at every point in a certain region. For example, we can speak of a temperature field in a room, as the temperature has a value at each point in the room. Because the electric field exerts a force, it is called a force field.

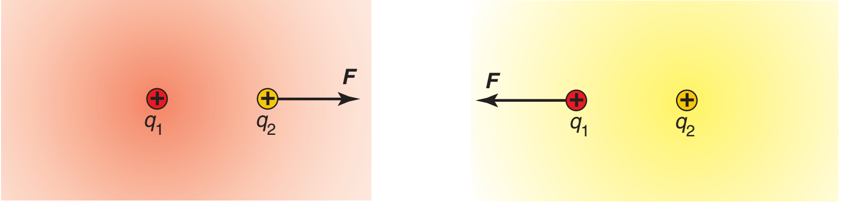
A different way of looking at the interaction between electric charges is in terms of **electric fields**. An electric field is a region where an electric charge experiences a force. On the field picture, every electric charge is surrounded by an electric field. If another charge, q, is placed in this electric field, the *field* exerts a force, F, on it. The field picture of electric force is illustrated in figure 6.17.

When two charges are brought close together, each charge is in the field of the other and experiences a force. This is illustrated in figure 6.18. Note that the field surrounding a charge does not exert any force on the charge itself, only on other charges placed in the field.



Field surrounding Q exerts a force on q.

Figure 6.17 ***Field picture of electric force***



Field surrounding q_1 exerts a force on q_2.

Field surrounding q_2 exerts a force on q_1.

Figure 6.18 ***Field picture of forces between two charges***

Electric field strength

An electric field exists at a point if an electric charge placed at the point experiences a force. *The direction of the electric field at a point is defined as the direction of the force that acts on a positive electric charge placed at the point.* A negative charge placed in an electric field experiences a force in the opposite direction to the field. This is illustrated in figure 6.19.

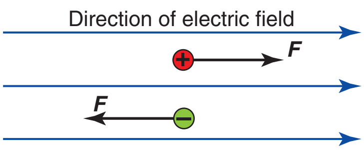


Figure 6.19 ***Direction of force on a charge placed in an electric field***

Compare an electric field with a gravitational field. An apple placed in the Earth's gravitational field experiences a force in the direction of the field, that is, downwards. The electrical situation is complicated by the fact that there are two different types of charge, positive and negative. A positive charge experiences a force in the direction of the field; a negative charge experiences a force in the opposite direction to the field. This is illustrated in figure 6.20.

The *magnitude* of the **electric field strength**, E, at a point is found by putting a charge, q, at the point and measuring the force, F, which the field exerts on it.

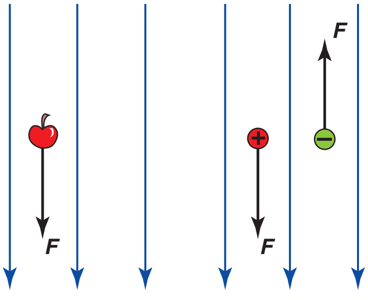


Figure 6.20 ***Comparison of electric and gravitational fields***

If a force of magnitude F acts on a charge, q, placed in an electric field, then the magnitude of the **electric field strength**, E, is given by the formula $E = \frac{F}{q}$.
The direction of the electric field strength is the direction of the force that acts on a positive charge placed in the field.

This is illustrated in figure 6.21. The magnitude of the electric field strength is defined by the formula $E = \frac{F}{q}$.

This can be expressed as: *The magnitude of the electric field strength at a point is the magnitude of the force per unit charge at the point.*

The SI unit of electric field strength is **newton coulomb^{-1}** (N C^{-1}).

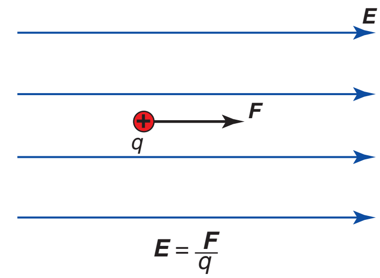


Figure 6.21 *Finding the magnitude of the electric field strength*

The **newton coulomb^{-1}** (N C^{-1}) is the unit of electric field strength.

SAMPLE PROBLEM 6.3

Calculating electric field strength

A charge of +3.0 μC, placed at a point in an electric field, experiences a force of 2.0 × 10^{-4} N east. Calculate the electric field strength at the point.

SOLUTION

The magnitude of the electric field strength is given by the formula:

$$
\begin{aligned}
E &= \frac{F}{q} \\
&= \frac{(2.0 \times 10^{-4})}{(3.0 \times 10^{-6})} \\
&= 6.7 \times 10^{1} \text{ N C}^{-1}.
\end{aligned}
$$

The positive charge experiences a force to the east, therefore the direction of the electric field is east.

SAMPLE PROBLEM 6.4

Calculating force on a charge in an electric field

A charge of −4.0 μC is placed at a point where the electric field strength is 6.0 × 10^{3} N C^{-1} north. Calculate the force that will act on the charge.

SOLUTION

$$
\begin{aligned}
E &= \frac{F}{q} \\
6.0 \times 10^{3} &= \frac{F}{(4.0 \times 10^{-6})} \\
F &= 2.4 \times 10^{-2} \text{ N}
\end{aligned}
$$

Note that the sign of q is not used in this calculation.

As the charge is negative, it will experience a force in the opposite direction to the field. Therefore, the direction of the force is south.

Force and electric field strength are vector quantities — they have magnitude and direction. We have defined the magnitude and the direction of the electric field strength separately, but it is possible to use a single definition that covers both.

If a force, $\boldsymbol{F}$, acts on a positive charge, $+q$, placed at a point in an electric field, then the electric field strength, $\boldsymbol{E}$, at the point is given by the equation: $\boldsymbol{E} = \frac{\boldsymbol{F}}{+q}$. That is, the magnitude and direction of the electric field strength at a point is equal to the magnitude and direction of the force per unit positive charge placed at the point.

Note the use of bold to indicate vector quantities. Vector quantities are studied later in the course.

Examples of electric fields

The **lines of electric field** are lines drawn on a diagram to represent the direction and magnitude of an electric field.

Electric fields are represented in a diagram by **lines of electric field**. These lines have no physical reality, but are very useful for picturing the direction and the magnitude of the electric field.

The *direction* of the electric field lines indicates the *direction* of the electric field.

The *spacing* of the electric field lines indicates the *magnitude* of the electric field. The closer together the lines, the stronger the field.

A point charge is a very small body with an electric charge.

Electric field surrounding a positive point charge

Consider a positive point charge, $+Q$. To determine the direction of the electric field surrounding $+Q$, a small positive test charge, $+q$, is placed near $+Q$. The direction of the force on $+q$ will be away from $+Q$, as shown in figure 6.22. The electric field surrounding $+Q$ will therefore point away from $+Q$.

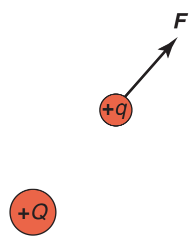


Figure 6.22 *Direction of the force on a positive test charge placed in the field surrounding a positive point charge*

Figure 6.23 shows the electric field surrounding a positive point charge, $+Q$. As you go further from the charge, the electric field lines are further apart indicating that the field is becoming weaker.

Figure 6.24 shows the forces that act on positive and negative charges placed in the field surrounding a positive point charge. A positive charge, $+q$, placed in the field will experience a force in the direction of the field; that is, away from the positive charge $+Q$. A negative charge, $-q$, placed in the field will experience a force in the opposite direction to the electric field; that is, towards the positive charge $+Q$.

In diagrams, electric fields are represented in two dimensions. It should be noted that electric fields are three dimensional.

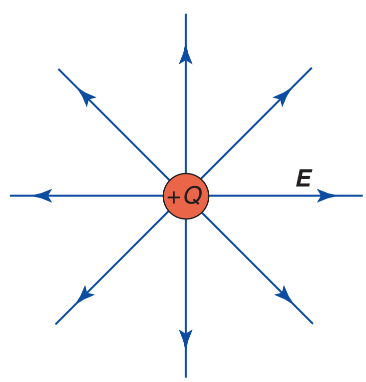


Figure 6.23 *Electric field surrounding a positive point charge*

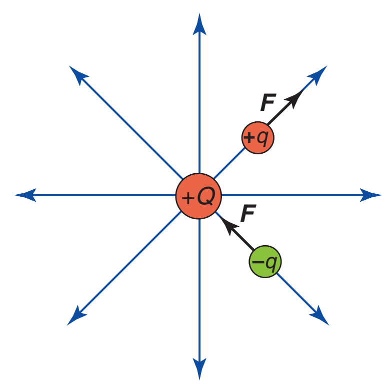


Figure 6.24 *Forces on charges placed in the field surrounding a positive point charge*

Electric field surrounding a negative point charge

Consider a negative point charge, $-Q$. To determine the direction of the electric field surrounding $-Q$, a small positive test charge, $+q$, is placed in the field. The force on $+q$ will be towards $-Q$. Therefore, the electric field surrounding $-Q$ will be towards $-Q$. This is illustrated in figure 6.25.

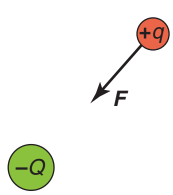


Figure 6.25 *Force on a positive test charge placed in the field surrounding a negative point charge*

Figure 6.26 (on page 101) shows the electric field surrounding a negative point charge, $-Q$.

Figure 6.27 (on page 101) shows the forces that act on charges placed in the field surrounding a negative point charge. A positive charge, $+q$, placed in the field will experience a force in the direction of the field; that is, towards the negative charge $-Q$. A negative charge, $-q$, placed in the field will experience a force in the opposite direction to the electric field; that is, away from the negative charge $-Q$.

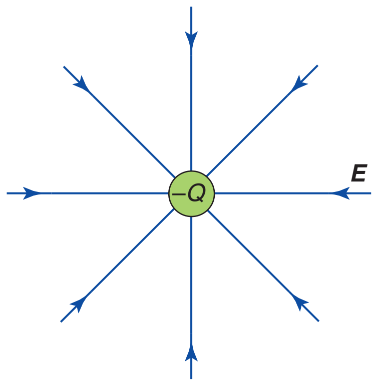


Figure 6.26 ***Electric field surrounding a negative point charge***

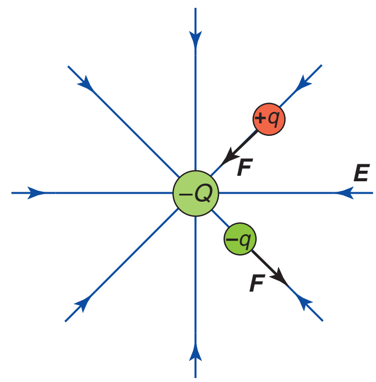


Figure 6.27 ***Forces on charges placed in a field surrounding a negative point charge***

A **uniform electric field** is an electric field with the same magnitude and direction at all points.

Uniform electric field

An electric field with the same magnitude and direction at all points is called a **uniform electric field**. If two parallel metal plates a short distance apart are given equal opposite charges, the field between the plates is uniform, as shown in figure 6.28. Note that, except at the edges of the plates:

- the lines of electrical field between the plates are parallel, showing that the electric field has the same direction at all points
- the lines of electrical field between the plates are equally spaced, showing that the electric field has the same magnitude at all points.

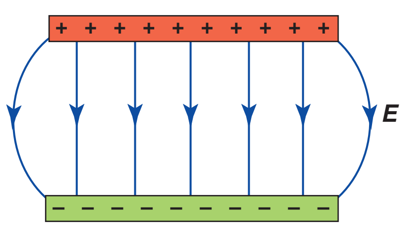


Figure 6.28 ***Uniform electric field***

Electric fields surrounding pairs of point charges

When there is more than one point charge producing an electric field, the fields from the individual charges combine to produce a single resultant field. Figures 6.29 to 6.33 show examples of electric fields produced by two point charges a small distance apart. In some of these examples, there are points where the fields from the two charges cancel one another. At these points, called *null points*, the electric field strength is zero. Null points are marked ‘n’ in the diagrams.

In drawing lines of electric field the following points should be noted:

- lines start on positive charges and end on negative charges
- lines never cross (the field cannot have two directions at a point)
- the greater the charge, the greater the number of lines starting or ending on it
- equal charges have equal numbers of lines starting or ending on them.

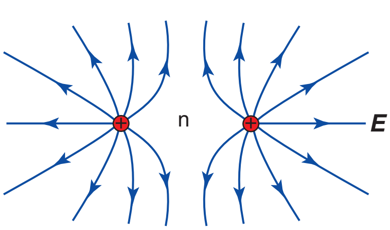


Figure 6.29 ***Electric field due to two equal positive point charges***

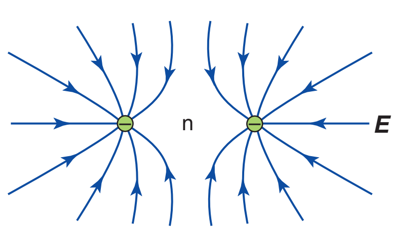


Figure 6.30 ***Electric field due to two equal negative point charges***

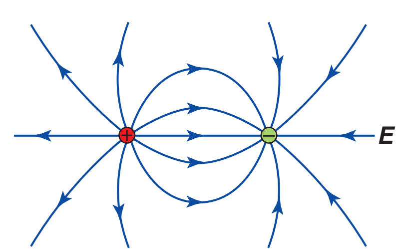


Figure 6.31 ***Electric field due to equal positive and negative point charges***

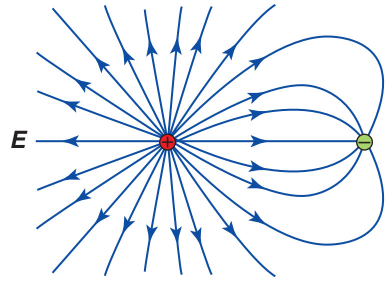


Figure 6.32 ***Electric field surrounding unequal positive and negative point charges***

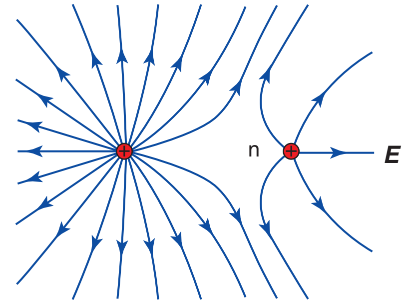


Figure 6.33 ***Electric field surrounding two unequal positive point charges***

6.3 ELECTRIC POTENTIAL ENERGY

Electric potential energy is the potential energy of an electric charge in an electric field.

Recall that potential energy is stored energy. An electric charge placed in an electric field has **electric potential energy**. The SI unit of electric potential energy is the joule (J).

Electric potential energy of a positive charge in an electric field

The electric potential energy of a charge in an electric field is similar to the gravitational potential energy of a mass in a gravitational field.

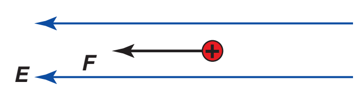
The free movement of a charge in an electric field is similar to a mass falling in the Earth's gravitational field.

Moving a positive charge in the opposite direction to an electric field is similar to raising a mass in a gravitational field.

A positive charge, $+q$, in an electric field will experience a force, $\boldsymbol{F}$, in the direction of the field. If the charge is free to move, it will move in the direction of the field, increasing in speed and therefore gaining kinetic energy. The gain in kinetic energy has come from a loss in electric potential energy. This is illustrated in figure 6.34.

To move the positive charge in the opposite direction to the field, energy must be expended to increase the electric potential energy of the charge.

When a positive charge moves in the direction of an electric field, its electric potential energy decreases. When it moves in the opposite direction to an electric field, its electric potential energy increases.

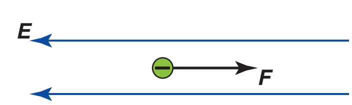


Charge accelerates in direction of field.

Figure 6.34 ***Positive charge moving freely in an electric field***

Electric potential energy of a negative charge in an electric field

A negative charge, $-q$, in an electric field will experience a force, $\boldsymbol{F}$, in the opposite direction to the field. If the negative charge is free to move, it will move in the opposite direction to the field, increasing in speed and therefore gaining kinetic energy. The gain in kinetic energy has come from a loss in electric potential energy. This is illustrated in figure 6.35.



Charge accelerates in opposite direction to field.

Figure 6.35 ***Negative charge moving freely in an electric field***

To move a negative charge in the direction of the field, energy must be expended to increase the electric potential energy of the charge.

When a negative charge moves in the opposite direction to an electric field, its electric potential energy decreases. When a negative charge moves in the direction of an electric field, its electric potential energy increases.

Potential difference

Electric **potential difference** between two points is the change in potential energy per unit charge moving between the two points.

The **potential difference** between two points in an electric field is the *change* in electric potential energy per coulomb of charge that moves between the points.

Consider a charge, q, that moves between two points in an electric field. If the change in the electric potential energy of the charge is W, then the potential difference, V, between the points is given by the formula: $V = \frac{W}{q}$.

Note that in this book the symbol 'V' is used for **potential difference** *and the symbol '*V*' for its unit, the* **volt**. *For example, $V = 10$ V. This distinction is not made in ordinary handwriting.*

Figure 6.36 shows a charge moving between two points in an electric field.

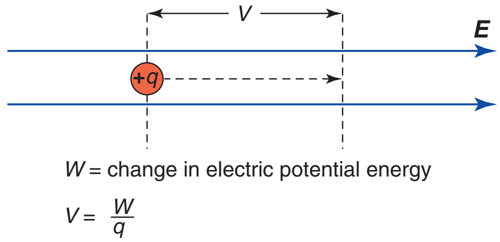


Figure 6.36 *Potential difference between two points in an electric field*

The **volt** (V) is the SI unit of potential difference.

Voltage is another name for potential difference.

The equation $V = \frac{W}{q}$ is not required by the syllabus. It is included here as a useful formula for making clear the meaning of potential difference. The SI unit of potential difference is the **volt** (V). A volt is equivalent to a joule coulomb^{-1}. Potential difference is also referred to as **voltage**.

SAMPLE PROBLEM 6.5

Potential difference

When a charge of 2.50×10^{-4} C moves between two points in an electric field, the electric potential energy of the charge changes by 5.00×10^{-2} J.

(a) Calculate the potential difference between the two points.

(b) Calculate the change in potential energy if a charge of 7.60×10^{-2} C were moved between the two points.

SOLUTION

(a)
$$V = \frac{W}{q}$$
$$= \frac{(5.00 \times 10^{-2})}{(2.50 \times 10^{-4})}$$
$$= 2.00 \times 10^{2} \text{ V}$$

(b)
$$V = \frac{W}{q}$$
$$2.00 \times 10^{2} = \frac{W}{(7.60 \times 10^{-2})}$$
$$W = (2.00 \times 10^{2}) \times (7.60 \times 10^{-2})$$
$$= 1.52 \times 10^{1} \text{ J}$$

A more complete treatment of potential difference distinguishes between a potential rise and a potential drop. This is similar to a rise in height and a drop in height in a gravitational field. Figure 6.37 illustrates the comparison between an electric field and a gravitational field.

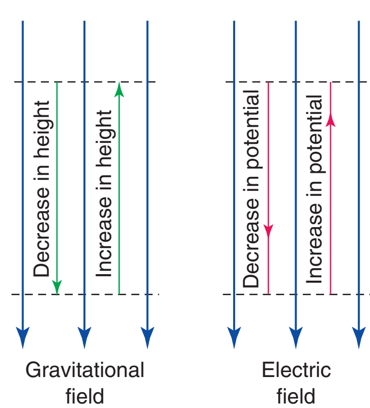


Figure 6.37 *Comparison between electric potential difference and change in height in a gravitational field*

6.4 ELECTRIC CURRENTS

A simple electric circuit

Before studying electric currents in more detail, we will look at a simple example to recall earlier work you have done on this subject. Figure 6.38 shows a familiar situation involving an electric current.

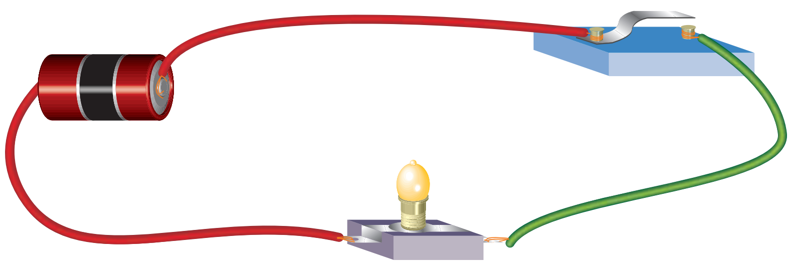


Figure 6.38 *A simple electric circuit*

Note the following:

- For the globe to light up there must be a complete conducting path between the terminals of the battery. The switch must be closed.
- The battery is necessary for a current to flow around the circuit. The ability of the battery to cause a current to flow is often referred to as its voltage.
- The battery has two terminals, marked positive and negative.
- The light globe resists the flow of the current. As a result of this resistance, the current causes the light globe to heat up to such an extent that it gives off light.
- The wires connecting the light globe to the battery do not heat up.

As an aid in representing electric circuits in diagrams, the symbols shown in figure 6.39 are used. Therefore, the circuit shown in figure 6.38 can be represented more simply as shown in figure 6.40.

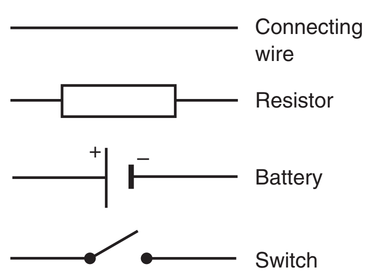


Figure 6.39 *Symbols for circuit components*

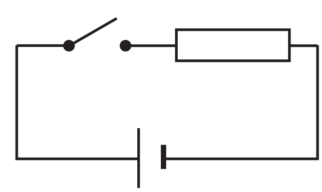


Figure 6.40 *Circuit diagram for circuit shown in figure 6.38*

eBookplus

eLesson:
The hydraulic model of current
eles-0029
Interactivity:
The hydraulic model of current
int-0053

What is an electric current?

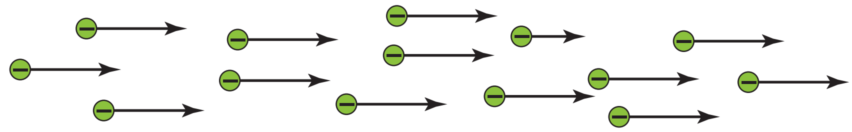
Electric current is the rate at which charge flows under the influence of an electric field.

Electric current is the rate at which charge flows under the influence of an electric field. The moving charges are called charge carriers.

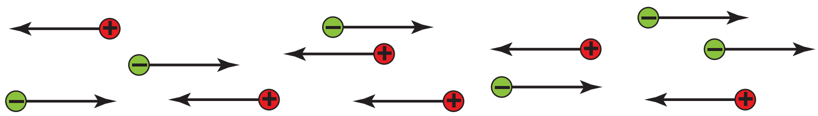
In a television tube, a beam of electrons travels along the tube to the screen. This is an electric current (see figure 6.41a on page 105).

When an electric current passes through a salt solution, positive and negative ions move in opposite directions through the solution (see figure 6.41b on page 105).

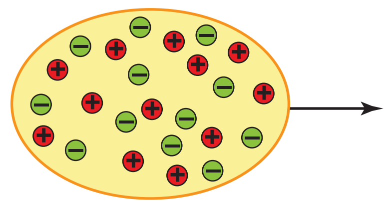
If a neutral body is moving, there are equal numbers of protons and electrons moving in the same direction at the same speed. There is no net movement of charge and therefore no current (see figure 6.41c).



(a) This is an electric current. A beam of electrons.



(b) This is an electric current. Movement of ions in a salt solution.



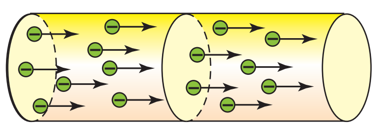
(c) This is not an electric current. Movement of a neutral body.

Figure 6.41 *Electric currents*

The ampere (A) is the SI unit of electric current. An ampere is equivalent to a coulomb second^{-1}.

The ampere is actually defined in terms of magnetic forces. This definition is beyond the scope of the preliminary course.

The SI unit of electric current is the **ampere** (A). (Ampere is usually abbreviated to amp.) An ampere is the amount of current flowing when a net charge of one coulomb flows through a cross-section of a conductor in one second. An ampere is equivalent to a coulomb second^{-1}. In terms of electrons, one ampere is the current that flows when 6.25×10^{18} electrons (approximately) pass through a cross-section of a conductor in one second (see figure 6.42).



When the current is 1 ampere, 6.25×10^{18} electrons pass through a cross-section of the conductor in one second.

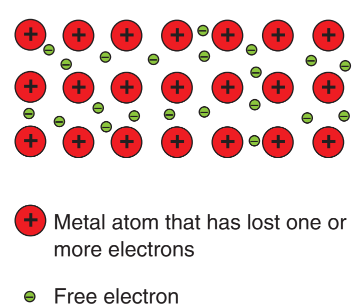
Figure 6.42 *Meaning of ampere in terms of electrons*

The symbol for current is *I*. For example, we write '$I = 5$ A'.

Conduction through a metal

Free electrons are electrons in a metal that are detached from their atoms and are free to move through the metal. A metal conducts an electric current by the movement of the free electrons.

In a metal, some electrons become detached from their atoms and are able to move freely within the metal. These are called **free electrons**. The atoms that have lost electrons become positively charged ions. The structure of a metal is illustrated in figure 6.43. The positively charged atoms form a lattice through which the free electrons move freely. The free electrons are the charge carriers that allow the metal to conduct an electric current.



Metal atom that has lost one or more electrons

Free electron
The free electrons are in constant random motion.

Figure 6.43 *Structure of a metal*

The free electrons in a metal are in constant random motion. Each free electron makes frequent collisions with positive ions of the lattice. At each collision the electron changes direction. Although the average speed of free electrons between collisions is of the order of 10^6 m s^{-1}, there is no net movement of the electrons, so there is no electric current. The random movement of a free electron through a metal lattice is shown in figure 6.44a.

If there is an electric field in a metal, the free electrons, being negatively charged, will experience a force in the opposite direction to the field. As a result of this force, there will be a net movement in the direction of the force superimposed on the random movement of the free electrons. This net movement is called **electron drift** and constitutes an electric current. Electron drift is illustrated in figure 6.44b. The drift velocity of an electron is of the order of 10^{-4} m s^{-1}, much smaller than the average speed of its random motion.

Electron drift is the slow movement of electrons through a conductor in the opposite direction to the electric field. This movement is superimposed on the much faster, random motion of the electrons.

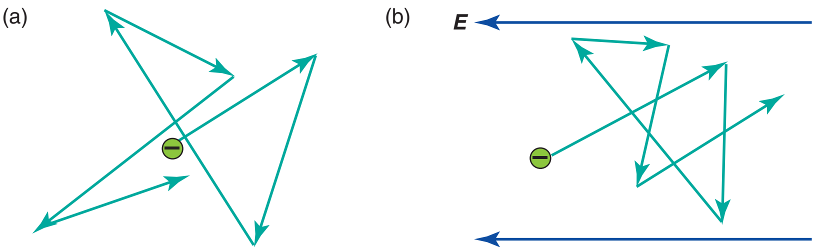


Figure 6.44 *Motion of free electron (a) with no electric field (b) with an electric field*

SAMPLE PROBLEM 6.6

Current in terms of electrons

A current of 5.00 A passes through a wire for 6.00 s. Calculate the number of electrons passing through a cross-section of the wire in this time.

SOLUTION

As the current is 5.00 A, a charge of 5.00 C will pass through a cross-section of the wire in 1 s. Therefore, in 6.00 s a charge of 6.00 × 5.00 C will pass through. 1 C of charge is equivalent to 6.25×10^{18} electrons. If n is the number of electrons, then:

$$\begin{aligned} n &= 6.00 \times 5.00 \times 6.25 \times 10^{18} \\ &= 1.88 \times 10^{20}. \end{aligned}$$

Energy transformation when a current passes through a metal

As the free electrons drift in the opposite direction to the field, they lose electric potential energy and gain kinetic energy. The free electrons continually collide with the positive ions in the metal lattice. In these collisions, the kinetic energy gained by the electrons is transferred to the positive ions of the lattice, causing them to vibrate with greater energy. The energy of vibration of the atoms of a body is heat energy. Thus, when an electric current flows through a metal, electric potential energy is transformed into heat energy.

Conductors and resistors

Connecting wires for electric circuits are made so that there is negligible conversion of electric potential energy into heat energy when a current passes through them. In studying electric circuits these components are referred to simply as conductors.

In other components in electric circuits, the conversion of electric potential energy into heat energy cannot be neglected. These components are referred to as **resistors**. The element of an electric heater is an example of a resistor. When an electric current flows through a resistor it loses electric potential energy, and heat energy is generated.

A **resistor** is a conductor in which the electric potential energy of a current is converted into heat energy.

Potential difference across a resistor

The **potential difference across a resistor** is the number of joules of electric potential energy dissipated for each coulomb of charge that passes through the resistor.

In a resistor, electric potential energy is converted into heat energy. The electric potential energy is said to be dissipated. The **potential difference across a resistor** is the number of joules of electric potential energy dissipated by each coulomb of charge that passes through the resistor.

Potential difference across a resistor is also referred to as the voltage across the resistor.

If potential energy, W, is dissipated when a charge, q, moves through the resistor, then the voltage, V, across the resistor is given by the formula:

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

For example, if there is a potential difference of 1000 V between the ends of the element in an electric heater, then for each coulomb that passes through the element, 1000 joules of electric potential energy is converted into heat energy.

SAMPLE PROBLEM 6.7

Calculation of heat energy generated when charge passes through a resistor

Calculate the number of joules of heat energy that are generated when 2.0×10^1 C of charge pass through a resistor with a potential difference of 6.0×10^1 V between the ends.

SOLUTION

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

$$6.0 \times 10^1 = \frac{W}{2.0 \times 10^1}$$

$$W = 1.2 \times 10^3 \text{ J}$$

As 1.2×10^3 J of electric potential energy have been dissipated, 1.2×10^3 J of heat energy will have been generated.

Power supply

The **power supply** is a source of electric potential energy.

When a current flows through a resistor, electric potential energy is transformed into heat energy. Therefore, to maintain a current, there must be a source of electric potential energy. The source of electric potential energy in a circuit is called the **power supply**.

Most power supplies convert another form of energy into electric potential energy. For example, a battery converts chemical energy into electric potential energy, while a photoelectric cell converts light energy into electric potential energy.

In the school laboratory, the power supply is usually a power pack that converts electric potential energy from the mains supply into a form of electric potential energy that is more suitable for school use. How this is done will be studied in the HSC course.

How a power supply produces a current

A power supply separates positive and negative charges to produce a positively charged terminal and a negatively charged terminal. If a conductor joins the positive and negative terminals of a power supply, an electric field is established through the conductor from the positive terminal to the negative terminal of the power supply. If the conductor is a metal, this field will exert forces on the free electrons in the opposite direction to the field, causing them to drift towards the positive terminal. This drift of free electrons is an electric current. This is illustrated in figure 6.45. When the conducting path is established, the current is produced almost simultaneously at all points around the circuit.

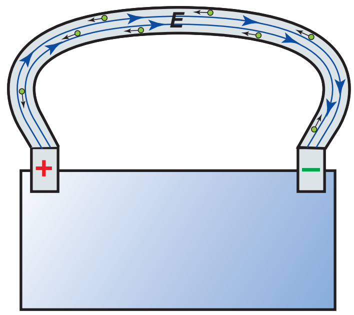


***Figure 6.45** Movement of free electrons in a metal joining the terminals of a power supply*

Potential difference across a power supply

Recall that a power supply gives electric potential energy to electric charge. The **potential difference across a power supply** is the number of joules of electric potential energy given to each coulomb of charge that passes through the power supply. This potential difference is often referred to as the voltage across the power supply.

The potential difference across a power supply is the number of joules of electric potential energy given to each coulomb of charge that passes through the power supply.

For example, a 5 volt battery gives 5 joules of electric potential energy to each coulomb of charge that passes through the battery.

If electric potential energy, W, is given to a charge, q, that passes through the power supply, then the potential difference, V, across the power supply is given by the formula:

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

SAMPLE PROBLEM 6.8

Calculation of electric energy generated when charge passes through power supply

Calculate the amount of electric potential energy given to 2.00 coulomb of charge that passes through a 1.50×10^2 V power supply.

SOLUTION

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

$$1.50 \times 10^2 = \frac{W}{2.00}$$

$$W = 3.00 \times 10^2 \text{ J}$$

Conventional current direction

It is now known that a current through a metal consists of a drift of negative charges from the negative to the positive terminal of a power supply (see figure 6.46a). Before this was known, the current was considered to be a flow of positive charge from the positive to the negative terminal of the power supply. This direction is still used today as the *conventional current direction* (see figure 6.46b). When the current direction is marked in a circuit diagram, the conventional current direction is always used.

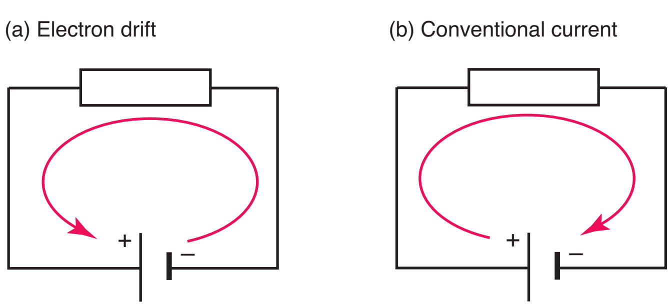


***Figure 6.46** (a) Electron drift direction (b) Conventional current direction*

Current at different points in a circuit

An **ammeter** is an instrument used to measure the electric current in an electric circuit. An ammeter is connected into a circuit in series.

An **ammeter** is an instrument used to measure current. Figure 6.47 shows an ammeter.

To measure the current through a point in a circuit, the ammeter must be inserted into the circuit in series as shown in figure 6.48. (Note the symbol for the ammeter.)

If the current is measured on each side of a resistor, it is found to have the same value on both sides. This is illustrated in figure 6.49. The current does not get less as it passes through the resistor. Current does not get 'used up'. What gets used up, as the current flows through a resistor, is electric potential energy.

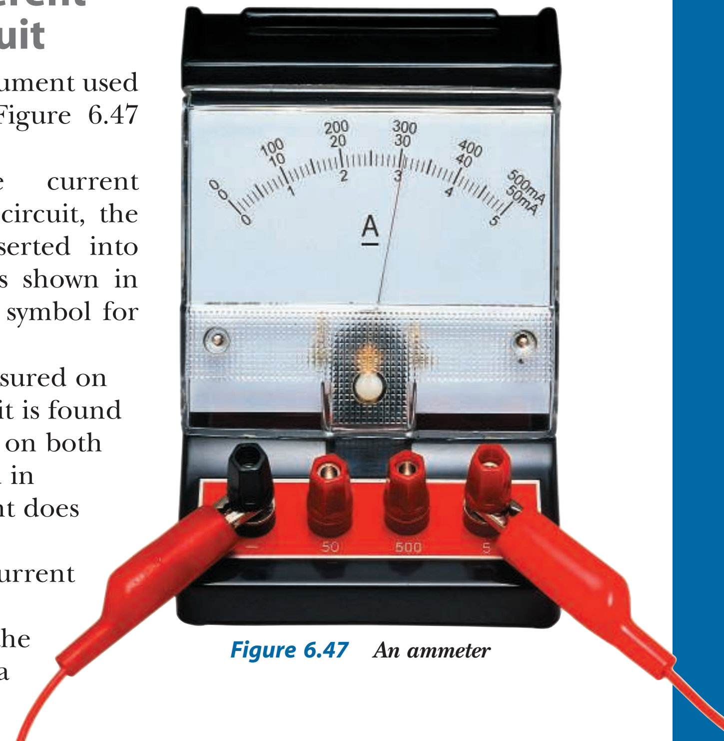


Figure 6.47 *An ammeter*

An ammeter has a very low resistance. When it is connected in series into a circuit it does not significantly change the currents and voltages in the circuit.

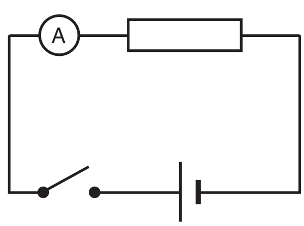
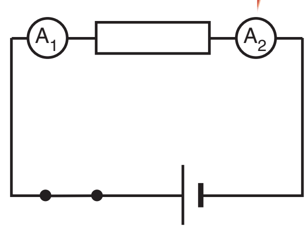


Figure 6.48 *Connecting an ammeter into a circuit*



Ammeters A_1 and A_2 will read the same.

Figure 6.49 *Current at different points of a circuit*

Potential differences across components of a circuit

A **voltmeter** is an instrument used to measure the potential difference across a component in an electric circuit. A voltmeter is connected into a circuit in parallel.

A **voltmeter** is used to measure the potential differences across components of a circuit.

Figure 6.50 shows a voltmeter. To measure the potential difference across a component of a circuit, the voltmeter is connected in parallel with the component. This is illustrated in figure 6.51. (Note the symbol for a voltmeter.)

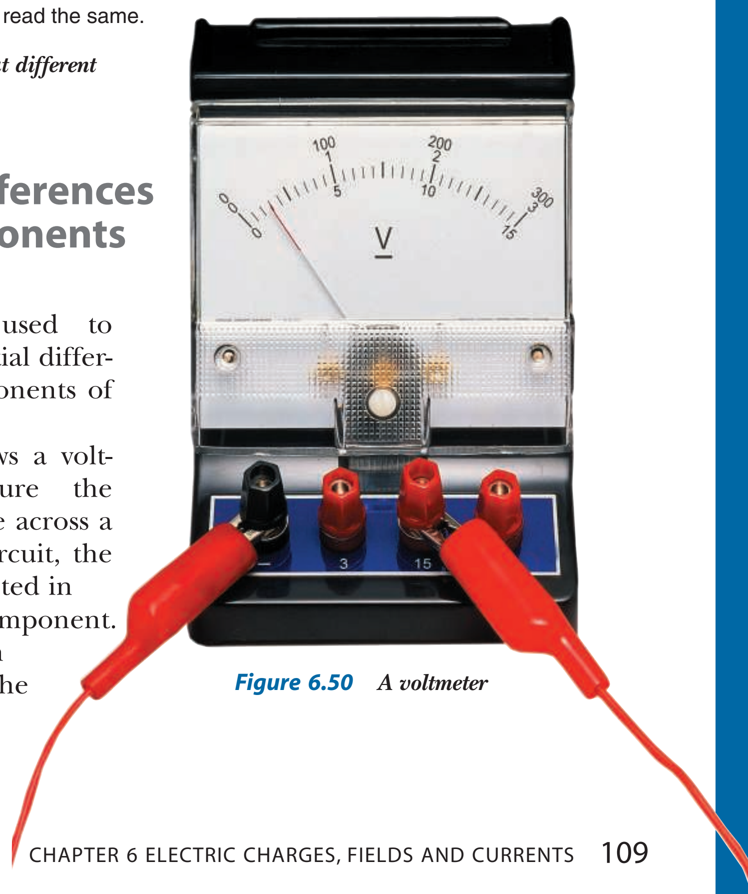


Figure 6.50 *A voltmeter*

A voltmeter has a very high resistance. When it is connected in parallel it does not significantly change the currents and voltages in the circuit.

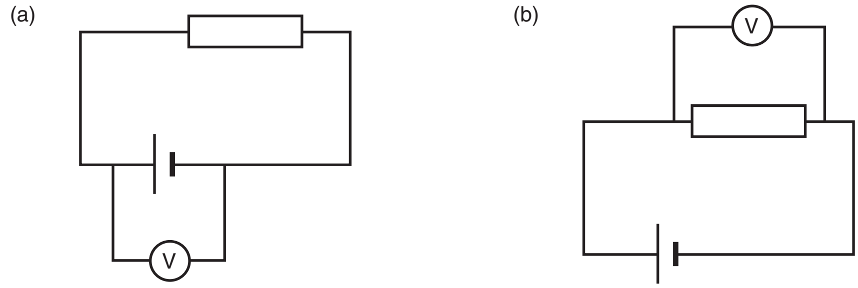
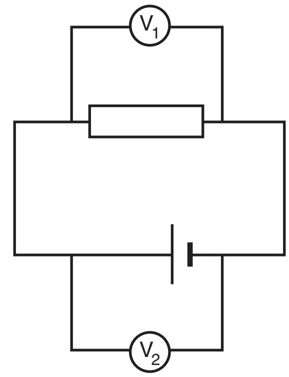


Figure 6.51 *Voltmeter measuring potential difference across (a) a power supply and (b) a resistor*

Consider the potential differences across the various components of the circuit shown in figure 6.52. The potential difference across each lead wire is zero. This is because there is negligible conversion of electric potential energy into heat energy in a good conductor. The potential difference across the power supply is equal to the potential difference across the resistor. This is because, for each coulomb of charge, the electric potential energy generated in the power supply is equal to the electric potential energy dissipated in the resistor.

The potential difference across a power supply is called a potential (or voltage) rise, as the current gains electric potential energy when it passes through a power supply. The potential difference across a resistor is called a potential (or voltage) drop, as the current loses electric potential energy as it passes through a resistor.



V_1 and V_2 will read the same.

Figure 6.52 *Potential differences around a circuit*

6.1

Potential difference in a simple circuit

The resistance of a resistor is the potential difference across the resistor divided by the current passing through the resistor.

$$R = \frac{V}{I}$$

This formula $R = \frac{V}{I}$ can also be written $V = IR$.

The ohm (Ω) is the unit of resistance.

Resistance

When a current, I, passes through a resistor, electric potential energy is dissipated, so there will be a potential drop, V, across the resistor. The **resistance**, R, of a resistor is defined by the formula: $R = \frac{V}{I}$.

This definition of resistance is illustrated in figure 6.53.

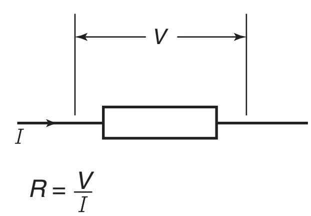


Figure 6.53 *Definition of resistance*

The SI unit of resistance is the **ohm** (Ω). If a current of 1 A is passing through a conductor of resistance 1 Ω, the potential difference between the ends of the conductor will be 1 V. The symbol for resistance is 'R'. For example, we write '$R = 5$ Ω'.

Resistance of a good conductor

A good conductor is one in which there is negligible conversion of electric potential energy into heat energy. Therefore, there will be negligible potential difference, V, across a good conductor. That is, $V = 0$.

The resistance of the conductor is given by the formula: $R = \frac{V}{I}$.

Since $V = 0$, $R = 0$. Therefore, a good conductor has zero (or negligible) resistance.

6.2

Measuring resistance

SAMPLE PROBLEM 6.9

Calculating resistance

When a current of 5.0×10^{-1} A is passing through a certain resistor, there is a potential difference of 2.0×10^{1} V between the ends of the resistor. Calculate the resistance of the resistor.

SOLUTION

$$\begin{aligned} R &= \frac{V}{I} \\ &= \frac{2.0 \times 10^{1}}{5.0 \times 10^{-1}} \\ &= 4.0 \times 10^{1}\ \Omega \end{aligned}$$

SAMPLE PROBLEM 6.10

Calculating potential difference

A current of 3.0 A flows through a 6.0 Ω resistor. Calculate the potential difference across the resistor.

SOLUTION

$$\begin{aligned} V &= IR \\ &= 3.0 \times 6.0 \\ &= 1.8 \times 10^{1}\ \text{V} \end{aligned}$$

SAMPLE PROBLEM 6.11

Current through, and voltage across, a resistor

When a current of 5.00 A passes through a resistor there is a potential difference of 1.00×10^{2} V across the resistor.

(a) Calculate the resistance.

(b) Calculate the current flowing if the potential drop across the resistor was 40.0 V.

(c) Calculate the potential drop across the resistor if the current was 10.0 A.

SOLUTION

(a)
$$\begin{aligned} V &= IR \\ 1.00 \times 10^{2} &= 5.00 \times R \\ R &= \frac{1.00 \times 10^{2}}{5.00} \\ &= 20.0\ \Omega \end{aligned}$$

(b)
$$\begin{aligned} V &= IR \\ 40.0 &= I \times 20.0 \\ I &= \frac{40.0}{20.0} \\ &= 2.00\ \text{A} \end{aligned}$$

(c)
$$\begin{aligned} V &= IR \\ &= 10.0 \times 20.0 \\ &= 2.00 \times 10^{2}\ \text{V} \end{aligned}$$

eBook*plus*

Weblink:
Ohm's Law applet

PHYSICS IN FOCUS

Ohm's Law

Ohm's Law states: *The potential drop across a resistor is proportional to the current passing through the resistor:* $V \propto I$.

Ohm's Law applies only to resistors with *constant resistance*; that is, to resistors whose resistance is the same no matter what current is passing through them. For such resistors: $\frac{V}{I} = R$ (where R is constant). Ohm's Law can be written: $V = IR$ (where R is constant).

Resistors that obey Ohm's Law are called ohmic resistors. For an ohmic resistor, the graph of V against I will be a straight line. The slope of the graph will equal the constant resistance. This is illustrated in figure 6.54a.

(continued)

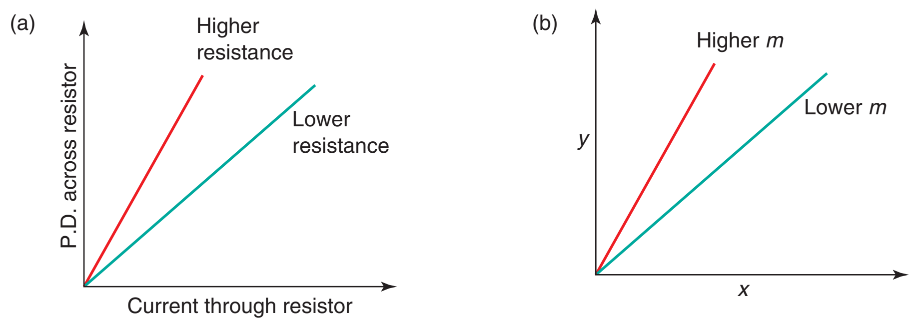


Figure 6.54 *(a) Graph of V against I for an ohmic resistor (b) Graph of $y = mx$ where m is constant*

The following are equivalent statements about a certain type of resistor:

- the resistor is an ohmic resistor
- the resistor obeys Ohm's Law
- the resistance of the resistor is constant
- the graph of V against I for the resistor is a straight line
- voltage is proportional to current for the resistor.

Compare the equation $V = IR$ where V and I vary and R remains constant with the familiar mathematical equation $y = mx$ for a straight line through the origin with gradient m. This is illustrated in figure 6.54b.

The resistance of a resistor is not always constant. For example, it is possible that a resistor may have a resistance of 2.00 Ω when the current is 1.00 A and a resistance of 2.20 Ω when the current is 2.00 A. Such a resistor does not obey Ohm's Law.

If a resistor does not obey Ohm's Law, it is called a non-ohmic resistor. One of the reasons that a resistor does not have constant resistance is that its temperature increases when higher currents pass through it. We will see later in this chapter that the resistance increases if the temperature rises.

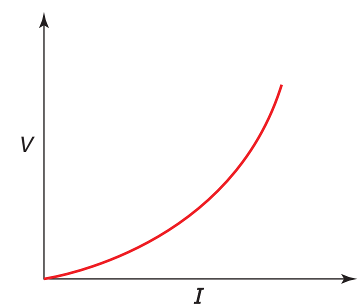


Figure 6.55 *V against I for a non-ohmic resistor*

For a non-ohmic resistor, the graph of V against I is not a straight line. Figure 6.55 shows a graph of V against I for a typical non-ohmic resistor.

The following are equivalent statements about this type of resistor:

- the resistor is non-ohmic
- the resistor does not obey Ohm's Law
- the resistance of the resistor is not constant
- the graph of V against I for the resistor is not a straight line.

PRACTICAL ACTIVITIES

6.3

Dependence of resistance on length of resistance wire

Factors that determine resistance

The resistance of a conductor is a result of collisions between the free electrons and the lattice of positive ions. The greater the number of collisions, the greater the resistance. The resistance of a particular conductor is determined by four factors:

- length
- area of cross-section
- material
- temperature.

Resistance and length

Consider free electrons drifting through a metal wire. The longer the wire, the greater the chance of a collision between a free electron and an ion in the lattice. Therefore, the longer the wire, the greater its resistance. If two conductors differ only in length, the longer conductor will have the greater resistance. The resistance, R, is proportional to the length, l: $R \propto l$.

If two conductors, differing only in length, have lengths l_1 and l_2 and resistances R_1 and R_2, then: $\frac{R_1}{R_2} = \frac{l_1}{l_2}$. If the length of a conductor is doubled, keeping all the other factors the same, the resistance will be doubled. This is illustrated in figure 6.56.

A useful comparison is with the flow of water through a pipe. It is more difficult for water to flow through a long pipe, than to flow through a short pipe.

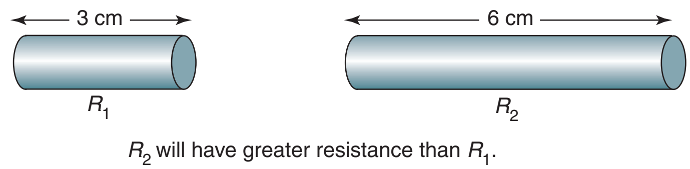


Figure 6.56 ***Dependence of resistance on length***

SAMPLE PROBLEM 6.12

Resistance and length

A 2.0 cm length of wire has a resistance of 1.6 Ω. What would be the resistance of 1.0×10^2 cm of the wire?

SOLUTION

The resistance of 1.0 cm of the wire is $\frac{1.6\ \Omega}{2.0\ \text{cm}}$.

Therefore, the resistance, R, of 1.0×10^2 cm of the wire will be:

$\left(\frac{1.6}{2.0}\right) \times (1.0 \times 10^2)\ \Omega$.

Therefore, $R = 8.0 \times 10^1\ \Omega$.

Resistance and area of cross-section

Consider free electrons drifting through a metal wire. The smaller the area of cross-section of the wire, the greater the chance of a collision of a free electron with an ion in the lattice. Therefore, the smaller the area of cross-section of the wire, the greater its resistance.

If two conductors differ only in area of cross-section, the conductor with the greater area of cross-section will have the lesser resistance. The resistance, R, of a wire is inversely proportional to the area of cross-section, A: $R \propto \frac{1}{A}$.

Note that doubling the area of cross-section is not the same as doubling the diameter of a wire. If the diameter is multiplied by two, the area of cross-section is multiplied by four.

If two conductors, differing only in area of cross-section, have areas of cross-section A_1 and A_2 and resistances R_1 and R_2, then: $\frac{R_1}{R_2} = \frac{A_2}{A_1}$.

If the area of cross-section of a conductor is doubled, keeping all the other factors the same, the resistance will be halved. This is illustrated in figure 6.57.

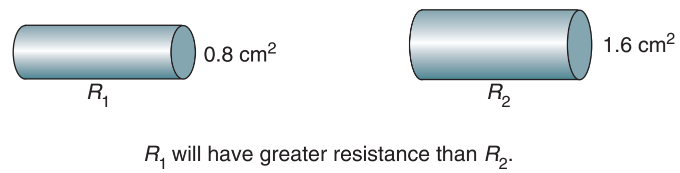


Figure 6.57 ***Dependence of resistance on area of cross-section***

The flow of water comparison applies here also. It is more difficult for water to flow through a narrow pipe than to flow through a wide pipe.

SAMPLE PROBLEM 6.13

Resistance and area of cross-section

Two pieces of resistance wire, X and Y, have the same length. Wire X has a cross-sectional area of 1.00 mm^2, and a resistance of 5.00 Ω. Wire Y has a cross-sectional area of 4.00 mm^2. What will be the resistance of wire Y?

SOLUTION

$$\frac{\text{Area of cross-section of X}}{\text{Area of cross-section of Y}} = \frac{1.00}{4.00}.$$

Therefore: $$\frac{\text{Resistance of X}}{\text{Resistance of Y}} = \frac{4.00}{1.00}.$$

Therefore: $$\frac{5.00}{\text{Resistance of Y}} = \frac{4.00}{1.00}.$$

Therefore: $$\text{Resistance of Y} = \frac{5.00 \times 1.00}{4.00}$$

$$= 1.25\ \Omega.$$

Resistance and material

When a free electron is drifting through a wire, the chance of a collision with an ion in the lattice depends, in a complex way, on what metal the wire is made of.

If two conductors have the same length and area of cross-section and are at the same temperature but are made of different materials, they will have different resistances.

We can compare a number of different materials. As copper is a very good conductor, we will compare the resistances of other materials with copper. Suppose we had a sample of copper with a resistance of 1.00×10^{-2} Ω, and identically shaped samples of other materials at the same temperature. Table 6.1 shows the resistances of the other materials.

Table 6.1 Comparative resistances of materials

MATERIAL	RESISTANCE (Ω)
Silver	0.94×10^{-2}
Copper	1.0×10^{-2}
Aluminium	1.6×10^{-2}
Iron	5.7×10^{-2}
Steel	11×10^{-2}
Constantan	26×10^{-2}
Nichrome	58×10^{-2}
Glass	approx. 10^{17}
Teflon	approx. 10^{21}

When a conductor with negligible resistance is required, copper is commonly used. When a conductor is required to have resistance, for example, in a heating coil, a material such as nichrome is used. Materials such as glass are used to make insulators.

For household circuits, copper wiring is used. Aluminium and steel (iron) are usually used for transmission lines as copper is too expensive and is not mechanically strong enough.

PHYSICS FACT

Benjamin Franklin originally made lightning conductors out of iron. When a lightning strike caused the iron in one of his lightning conductors to melt he started using copper.

Resistance and temperature

When the temperature of a conductor is increased, the ions in the lattice vibrate with greater amplitude. This increases the chance of a collision between a free electron and an ion in the lattice. Therefore, increasing the temperature of a conductor increases its resistance.

As an example, consider a conductor made of copper with a resistance of 1.000×10^{-3} Ω at 0°C. If its temperature is raised to 100°C, its resistance will be 1.393×10^{-3} Ω.

Direct and alternating voltages and currents

A school power supply (power pack) usually has two pairs of output terminals marked DC (direct current) and AC (alternating current). The DC terminals are marked positive and negative and are like the terminals of a battery. If a circuit is connected to the DC terminals, a current is produced in which the charge carriers move continuously in one direction.

Alternating voltages and currents are studied in more detail in chapter 7.

The polarity of a pair of terminals means which is positive and which is negative. The AC terminals periodically change their polarities — first one terminal being the positive terminal and then the other. The AC terminals provide an alternating voltage. If a circuit is connected to the AC terminals, a current will be produced in which the charge carriers move backwards and forwards periodically. Such a current is called an alternating current.

PHYSICS IN FOCUS

Resistance thermometers

The change of resistance of a conductor with temperature change can be used to make a thermometer. Such a thermometer can be used over a much greater range of temperatures than a liquid-in-glass thermometer.

The metal element of such a thermometer consists of a fine wire (approx. 10^{-4} m diameter). As this element is fragile, it is wound around a support made of mica, a mineral which is an insulator with a high melting point. The element is connected to an electrical circuit so that its resistance can be measured. To use the thermometer, the element is inserted into the place where the temperature is to be measured. The resistance of the element is measured using the electric circuit connected to it, and the temperature is calculated from the known temperature–resistance characteristics of the element.

The most common metals used to make resistance thermometers are platinum, nickel and copper. Figure 6.58 shows how the relative resistance of copper varies with temperature. The relative resistance of a metal at a particular temperature is the ratio of the resistance of the metal at that temperature to its resistance at 0°C.

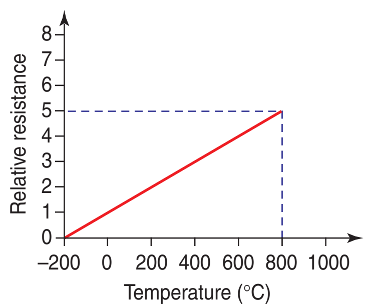


Figure 6.58 *Relative resistance against temperature for copper*

Copper wire elements are used between 120°C and 200°C; nickel elements between between 150°C and 300°C; platinum elements between 258°C and 900°C. A resistance thermometer can measure temperature to an accuracy of ± 0.01°C.

SUMMARY

- Protons have a positive electric charge; electrons have a negative electric charge.
- Like charges repel, unlike charges attract.
- The SI unit of charge is the coulomb (C).
- Positively charged bodies have a deficiency of electrons; negatively charged bodies have an excess of electrons.
- Bodies can be given electrostatic charges by friction, contact and induction.
- An electric field is a region where an electric charge experiences a force.
- Every electric charge is surrounded by an electric field.
- The electric field strength at a point is defined as the force per unit charge on a positive charge placed at the point: $\boldsymbol{E} = \frac{\boldsymbol{F}}{q}$.
- The direction of the electric field strength at a point is the direction of the force on a positive charge placed at the point.
- The direction of the electric field surrounding a positive charge is away from the charge. The direction of the electric field surrounding a negative charge is towards the charge.
- A charge in an electric field has electric potential energy.
- When a positive charge moves in the direction of an electric field, its electric potential energy decreases.
- When a negative charge moves in the opposite direction to an electric field its electric potential energy decreases.
- The potential difference between two points in an electric field is the change in electric potential energy per coulomb when a charge moves between the two points: $V = \frac{W}{q}$.
- The SI unit of potential difference is the volt. One volt is equivalent to one joule per coulomb.
- An electric current is a net movement of electric charge.
- The SI unit of current is the ampere (A). One ampere is equivalent to one coulomb per second.
- Metals are electrical conductors because they have free electrons that act as charge carriers.
- Insulators are materials that have no charge carriers and therefore, cannot carry an electric current.
- The potential difference (voltage) between the terminals of a power supply is the number of joules of electric potential energy given to each coulomb of electric charge.
- If a conductor connects the terminals of a power supply, a current will flow through the conductor. The movement of electrons is from the negative to the positive terminal of the power supply.
- The conventional current direction is from the positive to the negative terminal of the power supply.
- A resistor is a conductor that resists the movement of the current through it.
- When current flows through a resistor, electric potential energy is dissipated as heat energy.
- The potential difference (voltage) between the ends of a resistor is the number of joules of electric potential energy dissipated for each coulomb of charge that passes through the resistor.
- Resistance of a resistor is equal to the potential difference across the resistor divided by the current passing through the resistor: $R = \frac{V}{I}$.
- The SI unit of resistance is the ohm (Ω).
- The resistance of a resistor depends on length, cross-section area, material and temperature.
- In a direct current (DC) the charge carriers move continuously in one direction.
- In an alternating current (AC) the charge carriers move periodically backwards and forwards.
- An ammeter is used to measure current and is connected into a circuit in series.
- A voltmeter is used to measure electrical potential difference and is connected into a circuit in parallel.

QUESTIONS

1. When glass is rubbed with silk, the glass becomes positively charged. Explain, using a diagram, how this happens.
2. When referring to charged bodies, explain what is meant by:
   (a) excess of electrons
   (b) deficiency of electrons.
3. If one coulomb is equal to the charge on 6.25×10^{18} electrons, calculate the charge in coulombs on one electron.
4. A body has a positive charge of 2.00×10^{-6} coulombs. Calculate the number of electrons it has lost.
5. When a piece of perspex is rubbed with a piece of silk, 3.40×10^{5} electrons are transferred

CHAPTER REVIEW

from the perspex to the silk. Calculate the charge in coulombs on:
(a) the perspex
(b) the silk.

6. (a) Define the direction of an electric field.
(b) Using this definition, explain why the field surrounding a positive charge points away from the charge.

7. At a point in an electric field a positive charge of 3.00×10^{-6} C experiences a force of 6.50×10^{-4} N east. Calculate the electric field strength at the point in magnitude and direction.

8. At a point in an electric field, a negative charge of 2.50×10^{-5} C experiences a force of 7.50×10^{-6} N south. Calculate the electric field strength, in magnitude and direction, at the point.

9. At a certain point, the electric field strength is 4.30×10^{2} N C^{-1} east. Calculate the force, in magnitude and direction, on each of the following placed at the point:
(a) an electron
(b) a proton
(c) a charge of $+2.30 \times 10^{-4}$ C
(d) a charge of -6.50×10^{-4} C.

10. Figure 6.59 shows two points, X and Y, in an electric field.
(a) Explain why a positive charge would have more electric potential energy at X than at Y.
(b) Explain why a negative charge would have more electric potential energy at Y than at X.

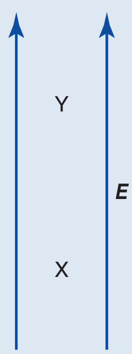


Figure 6.59

11. A positive charge, $+q$, is brought nearer to a positive charge, $+Q$. As $+q$ gets closer to $+Q$, discuss whether its electric potential energy increases or decreases.

12. A positive charge, $+q$, is brought closer to a negative charge, $-Q$. As the charge $+q$ gets closer to the charge $-Q$, discuss whether its electric potential energy increases or decreases.

13. When a charge of 3.75×10^{-4} C moves between two points in an electric field, the electric potential energy of the charge changes by 7.50×10^{-2} J.
(a) Calculate the potential difference between the two points.
(b) Calculate the change in potential energy if a charge of 2.60×10^{-3} C moved between the two points.

14. Define an electric current.

15. (a) Identify the charge carriers in a metal.
(b) Describe how these charge carriers move under the influence of an electric field in the metal.
(c) Describe how an electric field can be produced in the metal.

16. A current of 2.00 A is flowing through a wire.
(a) Calculate how many coulombs pass through a cross-section of the wire in 3.00 s.
(b) Calculate how many electrons pass through a cross-section of the wire in this time.

17. An electric current is flowing through a wire. A charge of 4.20 C passes through a cross-section of the wire in 3.00 s. Calculate the current in amps.

18. A current flows through a wire for 2.50 s. During this time 5.60×10^{18} electrons pass through a cross-section of the wire.
(a) Calculate the charge passing through a cross-section of the wire in this time.
(b) Calculate the current in amps.

19. A current of 5.00 amps is passing through a wire.
(a) Calculate the charge passing through a cross-section of the wire in:
(i) 1.00 second (ii) 1.00 minute.
(b) Calculate the number of electrons passing through a cross-section of the wire in each of these periods of time.

20. Explain how heat energy is produced when an electric current passes through a metal.

21. Explain what is meant by the potential difference across the ends of a resistor.

22. When 2.50 C pass through a certain resistor, 50.0 J of heat energy is generated. Calculate the potential difference across the resistor.

23. The potential difference across a certain resistor is 32.0 V. Calculate how much heat energy is produced when 12.0 C of electric charge pass through the resistor.

24. A battery is marked 25 V. Explain what this means.

25. Calculate the change in electric potential energy as 2.00 C of charge passes through a 2.50 V battery.

26. Explain what is meant by:
(a) a conductor
(b) a resistor
(c) an insulator.

27. When a current of 2.00 A passes through a certain resistor, there is a potential difference of 16.0 volts across it. Calculate the resistance of the resistor.

28. A coil of wire has a resistance of 3.20 Ω. Calculate the potential difference across the coil when there is a current of 2.00 A passing through it.

29. The following table refers to the potential difference, V, across a resistor of resistance R when a current, I, passes through it. Calculate the values of the missing quantities and complete the table.

	***V* (V)**	***I* (A)**	***R* (Ω)**
(a)	10.0	2.00	
(b)		6.00	12.0
(c)	16.0		32.0
(d)	3.00	18.0	
(e)	24.0		8.00
(f)		5.00	2.00

30. The following table refers to the change in potential energy, W, when an electric charge, q, passes through an electrical potential difference, V. Calculate the values of the missing quantities and complete the table.

	***V* (V)**	***q* (C)**	***W* (J)**
(a)	12.0	2.00	
(b)		25.0	50.0
(c)	6.00		42.0
(d)		8.00	32.0
(e)	4.00		0.50
(f)	2.00	16.0	

31. You are given four pieces of wire made of the same material. The lengths and diameters of the wires are given in the following table. List these in order of increasing resistance. Justify your answer.

	Length (cm)	**Diameter (mm)**
(a)	10	1
(b)	10	2
(c)	20	0.5
(d)	20	1

32. Demonstrate by means of diagrams the correct way of connecting into a circuit:
   (a) an ammeter
   (a) a voltmeter.

Note on practical activities

Instructions for activities involving electric currents:

- To measure the current through a component, connect an ammeter in series with the component.
- To measure the voltage across a component, connect a voltmeter in parallel with the component.
- Ammeters and voltmeters should be connected so that the connection from the positive terminal of the power supply goes to the positive terminal of the ammeter or voltmeter (see figure 6.60).

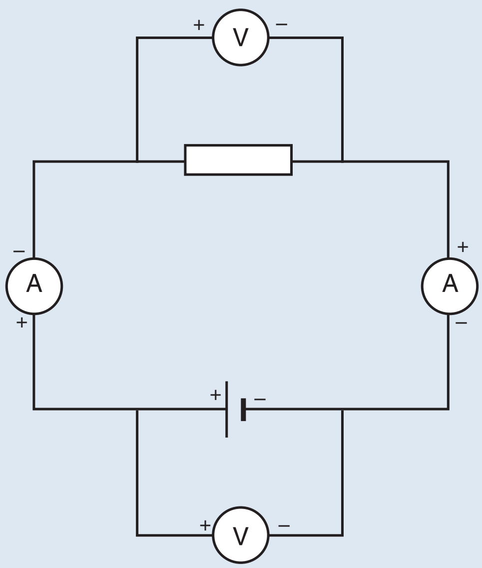


Figure 6.60 *Connecting ammeters and voltmeters to a power supply*

- If there is more than one range on a meter, start with the largest range (for example, if there is a 0–1 A range and a 0–5 A range on an ammeter, start with the 0–5 A range).
- Start with the lowest voltage setting on the power supply.
- If using a variable resistor to vary the current, start with its highest resistance.
- Tap the switch closed to check that the ammeters and voltmeters are correctly connected.
- When making a measurement, close the switch just long enough to make the measurement.
- Ask your teacher to check your circuit before switching on the power supply.
- Your teacher will advise you about what maximum current or maximum voltage you should use in a particular experiment.

6.1 POTENTIAL DIFFERENCE IN A SIMPLE CIRCUIT

Aim

To investigate the potential difference between different points around a simple circuit

Apparatus

power supply
voltmeter
two resistors, R_1 and R_2, of different resistance
switch
connecting wires

Theory

Potential difference is the change in electric potential energy per coulomb of charge. When a current flows through a circuit, several changes in potential occur. In the power supply there is an increase in electric potential energy and therefore, a potential rise. In a resistor, there is a decrease of electric potential energy and therefore, a potential drop. In a connecting wire, there is no change of electric potential energy and therefore, no potential difference.

Method

1. Connect the apparatus as shown in figure 6.61. Use R_1.

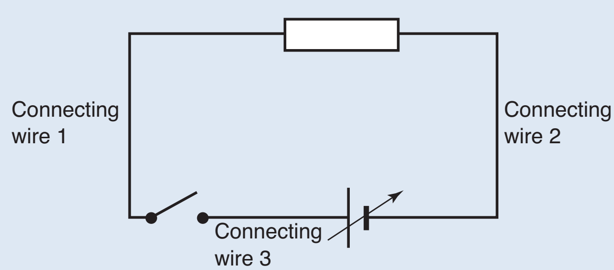


Figure 6.61

2. Set the voltage of the power supply to a low voltage.
3. Connect the voltmeter, in turn, across each of the following components shown in figure 6.61:
   - power supply
   - resistor
   - connecting wire 1

- connecting wire 2
- connecting wire 3 (including switch).

For each component, measure the voltage with the switch closed.

4. Repeat 3 using a higher voltage power supply.
5. Repeat 2, 3, and 4 using R_2.

Results

Record the results in a table as shown below.

R_1, low power supply voltage

COMPONENT	POTENTIAL DIFFERENCE (V)
Power supply	
Resistor	
Connecting wire 1	
Connecting wire 2	
Connecting wire 3	

Make similar tables for:

(i) R_1, high power supply voltage
(ii) R_2, low power supply voltage
(iii) R_2, high power supply voltage.

Analysis

1. What are the voltages across the connecting wires?
2. What is the relation between the voltage across the power supply and the voltage across the resistor?

Question

Explain, in terms of energy, why the potential rise across the power supply equals the potential drop across the resistor.

6.2 MEASURING RESISTANCE

Aim

To show the relationship between voltage and current, and measure the resistance of a resistor

Apparatus

power supply
ammeter
voltmeter
variable resistor
two resistors, R_1 and R_2, of different resistance
connecting wires
switch

Theory

The voltage drop, V, across a resistor is proportional to the current, I, through the resistor. The graph of V against I will be a straight line. The slope of the line will equal the resistance.

Method

1. Connect the apparatus as in the circuit diagram in figure 6.62, using R_1.

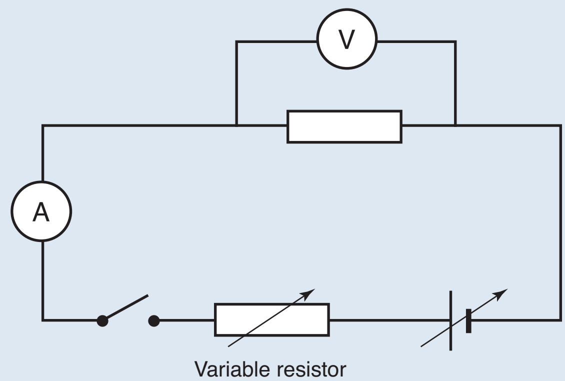


Figure 6.62

2. Pass at least five different currents through the resistor by varying both the voltage of the power supply and the resistance of the variable resistor.
3. For each value of the current, record the current through R_1 and the voltage across R_1.
4. Repeat steps 1, 2 and 3 using resistor R_2.

Results

Record your results in a table as shown below.

Measurements for R_1

V (V)	I (A)

Analysis

1. For each resistor, draw a graph of V against I using the same sheet of graph paper.
2. Measure the slope of each graph.
3. Use the slopes of the graphs to find the values of the resistance.

Questions

1. What are the shapes of the graphs?
2. Which graph has the steeper slope?
3. How do the values you found for the resistances compare with the values marked on the resistors?
4. Why would you expect your graphs to pass through the origin?
5. If your graphs do not pass through the origin suggest a reason.

6.3 DEPENDENCE OF RESISTANCE ON LENGTH OF RESISTANCE WIRE

Aim

To investigate how the resistance of a resistance wire varies with length

Apparatus

power supply
voltmeter
ammeter
variable resistor
switch
two different lengths of resistance wire, l_1 and l_2
gas jar
metre rule

Theory

The resistance of a conductor of constant thickness increases as its length increases. Resistance can be calculated by the formula: $R = \frac{V}{I}$.

Method

As the resistance of a wire varies with temperature, it is necessary to keep the resistance wire at constant temperature. This is done by immersing the resistance wire in water as shown in figure 6.63.

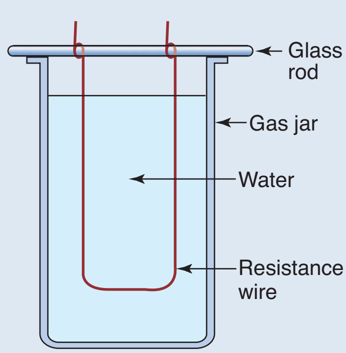


Figure 6.63 *Set-up for resistance wire*

1. Measure the lengths of resistance wire.
2. Connect the apparatus as shown in figure 6.64 using one of the resistance wires.

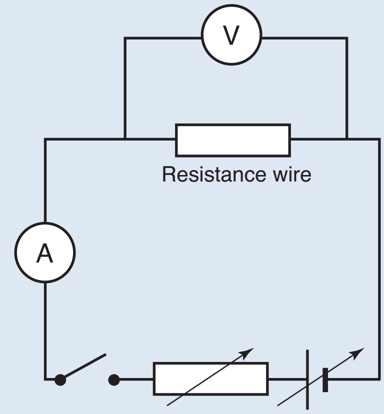


Figure 6.64

3. Vary the voltage of the power supply and the resistance of the variable resistor until the current has the maximum value given by your teacher.
4. Measure the voltage and the current with the switch closed.
5. Repeat using the other length of resistance wire.

Results

Record your results in a table as shown below and calculate the resistances.

LENGTH OF WIRE (cm)	CURRENT (A)	VOLTAGE (V)	RESISTANCE (Ω) $R = \frac{V}{I}$

Analysis

1. What is the ratio of the lengths of the wires?
2. What is the ratio of the resistances of the wires?
3. What conclusion can you draw from these values?

Questions

1. What would be the resistance of a 30 cm length of this wire?
2. What would be the resistance of a 2 cm length of this wire?
3. What is the resistance per unit length of this wire?

CHAPTER 7 THE HOUSEHOLD ELECTRICITY SUPPLY

Figure 7.1 *The electricity supply is delivered to houses through overhead or underground cables.*

Remember

Before beginning this chapter, you should be able to:

- design, draw and construct circuits with a number of components
- compare the advantages and disadvantages of series and parallel circuits.

Key content

At the end of this chapter you should be able to:

- identify the difference between series and parallel circuits
- compare series and parallel circuits in terms of voltage across components and current through them
- explain why there are different circuits for lighting, heating and other appliances in a house
- discuss the dangers of an electric shock to the muscles of the body from a 240 V AC mains supply and from DC voltages from appliances
- describe the functions of circuit breakers, fuses, earthing, double insulation and other safety devices in the home.

7.1 SERIES AND PARALLEL CIRCUITS

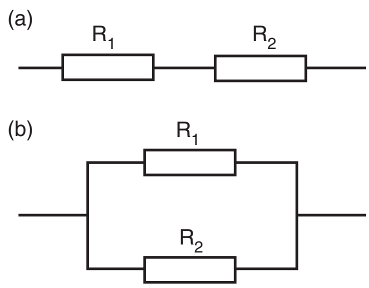


Figure 7.2 *(a) Series and (b) parallel connections*

When we deal with circuits with more than one resistor, it is convenient to label the resistors R_1, R_2, and so on. The resistances of the resistors are labelled R_1, R_2 and so on. (Note the use of italic to denote a variable, i.e. resistance.) Currents through the resistors are labeled I_1, I_2, and voltages across the resistors are labelled V_1, V_2, and so on. When it is not known how many resistors there are in a circuit, they are labelled R_1, R_2 . . . R_n.

Recall that resistors can be connected to a power supply in series or in parallel. Figure 7.2 shows two resistors, R_1 and R_2, connected (a) in series and (b) in parallel.

When resistors are connected in series, the same current passes through each resistor. When resistors are connected in parallel, the current divides and passes partly through each resistor.

PRACTICAL ACTIVITIES

7.1(a)

Current in a series circuit

Current in circuits

Current in a series circuit

Figure 7.3 shows three resistors, R_1, R_2 and R_3, connected in series. Note that the same current, I, flows through the power supply and through each resistor.

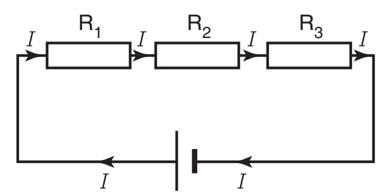


Figure 7.3 *Current in a series circuit*

SAMPLE PROBLEM 7.1

Current in a series circuit

In the circuit shown in figure 7.4, the current through the 8.0 Ω resistor is 2.0 A. Calculate the current through:

(a) the 5.0 Ω resistor

(b) the power supply.

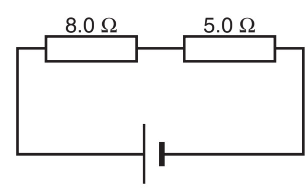


Figure 7.4

SOLUTION

As the resistors are connected in series to the power supply, the same current will flow through each resistor and through the power supply.

Therefore: (a) current through the 5.0 Ω resistor = 2.0 A

(b) current through the power supply = 2.0 A.

Current in a parallel circuit

In a parallel circuit, the current through the power supply divides, with part flowing through each resistor. Figure 7.5 shows three resistors, R_1, R_2 and R_3, connected in parallel. The current through the power supply, I, equals the sum of the currents in the resistors.

$$I = I_1 + I_2 + I_3$$

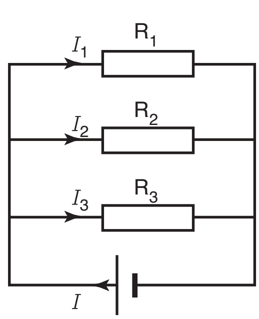


Figure 7.5 *Current in a parallel circuit*

PRACTICAL ACTIVITIES

7.1(b)

Current in a parallel circuit

If resistors with equal resistance are connected in parallel, the current will divide equally between them. If resistors with unequal resistance are connected in parallel, the current will divide unequally, with the greater current flowing through the resistor with the lesser resistance.

SAMPLE PROBLEM **7.2**

Current through parallel resistors

Consider the circuit shown in figure 7.6.

(a) Calculate the value of the current through R_2.

(b) Deduce which resistor has the greater resistance.

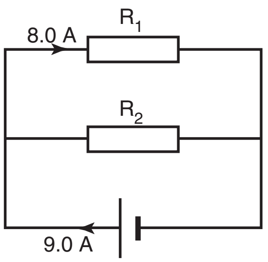


Figure 7.6

SOLUTION

(a) The resistors are in parallel with the power supply, therefore:

$I = I_1 + I_2$

$9.0 = 8.0 + I_2$.

Therefore: $I_2 = 1.0$ A.

(b) The resistor with the greater resistance has the smaller current. Therefore, R_2 has the greater resistance.

SAMPLE PROBLEM **7.3**

Current through equal, parallel resistors

Referring to the circuit in figure 7.7, calculate (a) the current through R_2 and (b) the current through the power supply.

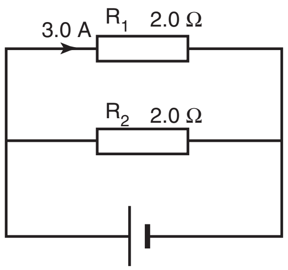


Figure 7.7

SOLUTION

(a) Since the two resistors in parallel have equal resistance, they will have equal current passing through them. Therefore,

$$I_2 = 3.0 \text{ A}.$$

(b) The current through the power supply is given by the formula:

$I = I_1 + I_2$

$= 3.0 + 3.0$

$= 6.0$ A.

Voltage in circuits

Recall from the previous chapter:

- the voltage *rise* in a power supply is equal to the electric potential energy given to each coulomb of charge that passes through the power supply
- the voltage *drop* in a resistor is equal to the electric potential energy lost by each coulomb of charge that passes through the resistor.

Voltage in a series circuit

Recall from the previous chapter that there is no voltage rise or fall across a connecting wire.

Figure 7.8 shows three resistors, R_1, R_2 and R_3, connected in series.

V_{ps} is the voltage rise in the power supply. V_1, V_2 and V_3 are the voltage drops in the resistors, R_1, R_2 and R_3.

As the current flows around the circuit, the potential energy gained by each coulomb of charge in the power supply equals the sum of the potential energies lost by each coulomb of charge in the resistors. Therefore,

$$V_{ps} = V_1 + V_2 + V_3.$$

If the resistors in series have equal resistance, the voltage drop across each will be the same. If the resistors in series do not have equal resistance, the resistor with the greater resistance will have the greater voltage drop across it.

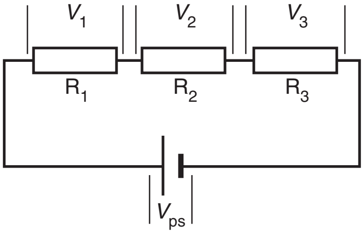
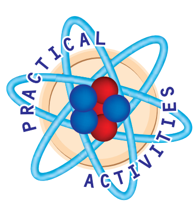


Figure 7.8 *Voltage in a series circuit*



7.1(c)

Voltage in a series circuit

SAMPLE PROBLEM 7.4

Voltage across series resistors

Consider in the circuit shown in figure 7.9.

(a) Calculate the voltage drop in the resistor R_2.

(b) Deduce which resistor has the greater resistance.

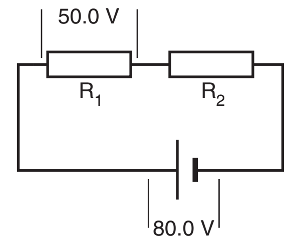


Figure 7.9

SOLUTION

(a) The two resistors are in series with the power supply. Therefore:

$V_{ps} = V_1 + V_2$

$80.0 = 50.0 + V_2$.

So, $V_2 = 30.0$ V.

(b) The greater resistance will have the greater voltage drop across it, therefore R_1 has the greater resistance.

SAMPLE PROBLEM 7.5

Voltage across equal series resistors

For the circuit shown in figure 7.10, calculate:

(a) the voltage drop across R_2

(b) the voltage gain across the power supply.

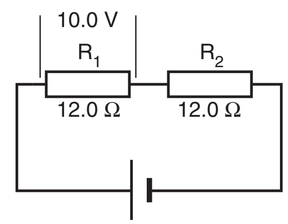


Figure 7.10

SOLUTION

(a) As R_1 and R_2 are equal and connected in series, there will be equal voltage drops across them. Therefore, the voltage drop across R_2 is 10.0 V.

(b) $V_{ps} = V_1 + V_2$

$= 10.0 + 10.0$

$= 20.0$ V

7.1(d)

Voltage in a parallel circuit

Voltage in a parallel circuit

Figure 7.11 shows three resistors, R_1, R_2 and R_3, connected in parallel. The voltage drops, V_1, V_2, V_3, across the resistors are *each* equal to the voltage rise, V_{ps}, across the power supply. To understand this, consider that the potential energy gained by each coulomb of charge in the power supply equals the total potential energy lost by one coulomb of charge in the resistors. As the coulomb of charge passes partly through each resistor, the voltage drop in each resistor must equal the voltage gain in the power supply. That is, $V_{ps} = V_1 = V_2 = V_3$.

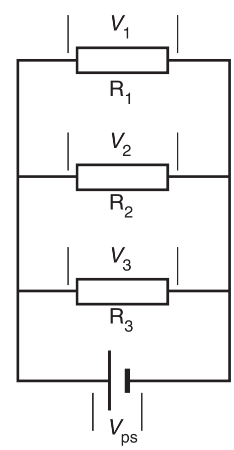


Figure 7.11 *Voltage in a parallel circuit*

SAMPLE PROBLEM 7.6

Voltage across parallel resistors

Referring to the circuit shown in figure 7.12, calculate:

(a) the voltage drop across the 24.0 Ω resistor

(b) the voltage gain across the power supply.

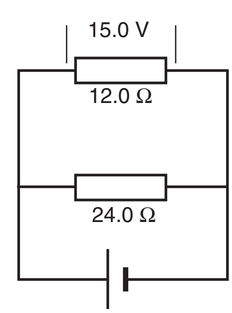


Figure 7.12

SOLUTION

(a) Since the resistors are connected in parallel with the power supply, the voltage drop is the same across each of them. Therefore, the voltage drop across the 24.0 Ω resistor = 15.0 V.

(b) The voltage rise across the power supply equals the voltage drop across each resistor. Therefore, $V_{ps} = 15.0$ V.

Analysing more complex series and parallel circuits

SAMPLE PROBLEM 7.7

Current and voltage in series circuit

Consider the circuit shown in figure 7.13.

(a) Calculate the current through the 2.50 Ω resistor.

(b) Calculate the current through the 6.00 Ω resistor.

(c) Calculate the voltage drop across the 6.00 Ω resistor.

(d) Calculate the voltage gain across the power supply.

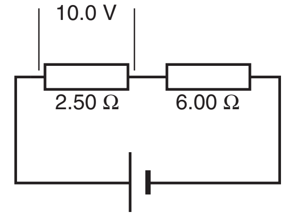


Figure 7.13

SOLUTION

(a) For the 2.50 Ω resistor:

$$\begin{aligned} I &= \frac{V}{R} \\ &= \frac{10.0}{2.50} \\ &= 4.00 \text{ A.} \end{aligned}$$

(b) Since the two resistors are in series, the same current flows through each. Therefore, the current through the 6.00 Ω resistor = 4.00 A.

(c) For the 6.00 Ω resistor:

$$\begin{aligned} V &= IR \\ &= 4.00 \times 6.00 \\ &= 24.0 \text{ V.} \end{aligned}$$

(d) As the resistors are in series with the power supply:

$$\begin{aligned} V_{ps} &= V_1 + V_2 \\ &= 10.0 + 24.0 \\ &= 34.0 \text{ V.} \end{aligned}$$

SAMPLE PROBLEM 7.8

Current and voltage in series circuit

For the circuit shown in figure 7.14, calculate:

(a) the current in the circuit

(b) the voltage drop across the 20.0 Ω resistor

(c) the voltage drop across the 15.0 Ω resistor.

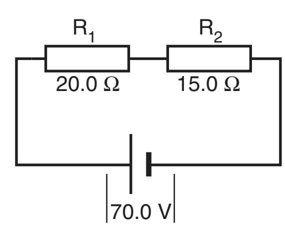


Figure 7.14

SOLUTION

(a) Let I be the current in the circuit.

$$\begin{aligned} V_1 &= IR_1 \\ &= I \times 20.0 \\ V_2 &= IR_2 \\ &= I \times 15.0. \end{aligned}$$

As the resistors are in series with the power supply,

$$\begin{aligned} V_{ps} &= V_1 + V_2 \\ 70.0 &= I \times 20.0 + I \times 15.0 \\ &= I \times 35.0 \\ I &= 2.00 \text{ A.} \end{aligned}$$

(b) $$\begin{aligned} V_1 &= IR_1 \\ &= 2.00 \times 20.0 \\ &= 40.0 \text{ V} \end{aligned}$$

(c) $$\begin{aligned} V_2 &= IR_2 \\ &= 2.00 \times 15.0 \\ &= 30.0 \text{ V} \end{aligned}$$

SAMPLE PROBLEM 7.9

Current and voltage in parallel circuit

In the circuit shown in figure 7.15, calculate:
(a) the current through the 4.00 Ω resistor
(b) the current through the 8.00 Ω resistor
(c) the current through the power supply.

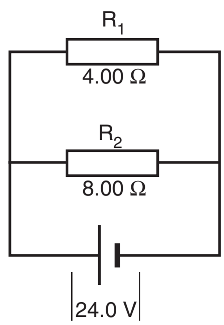


Figure 7.15

SOLUTION

(a) Since the resistors are in parallel with the power supply:

$$V_1 = V_{ps} = 24.0 \text{ V.}$$

For the 4.00 Ω resistor:

$$I_1 = \frac{V_1}{R_1} = \frac{24.00}{4.00} = 6.00 \text{ A.}$$

(b) By similar reasoning applied to the 8.00 Ω resistor:

$$I_2 = 3.00 \text{ A.}$$

(c) As the resistors are in parallel with the power supply:

$$I = I_1 + I_2 = 6.00 + 3.00 = 9.00 \text{ A.}$$

SAMPLE PROBLEM 7.10

Current and voltage in parallel circuit

In the circuit shown in figure 7.16, calculate:
(a) the voltage drop across the 10.0 Ω resistor
(b) the voltage drop across the 20.0 Ω resistor
(c) the current through the 20.0 Ω resistor
(d) the current through the power supply.

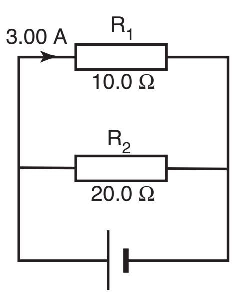


Figure 7.16

SOLUTION

(a) For the 10.0 Ω resistor:

$$V_1 = I_1R_1 = 3.00 \times 10.0 = 30.0 \text{ V.}$$

(b) As the resistors are in parallel there will be the same voltage drop across each.
Therefore, the voltage drop across the 20.0 Ω resistor is 30.0 V.

(c) For the 20.0 Ω resistor:

$$V_2 = I_2R_2$$

$$30.0 = I_2 \times 20.0.$$

Therefore, $I_2 = 1.50$ A.

(d) The current through the power supply is given by:

$$I = I_1 + I_2 = 3.00 + 1.50 = 4.50 \text{ A.}$$

SAMPLE PROBLEM 7.11

Number of resistors in a series circuit

Calculate how many 2.00 Ω resistors, connected in series to a 24.0 V power supply, would result in a current of 4.00 A through the circuit?

SOLUTION

Let the number of resistors connected in series be N. The sum of the voltage drops across the resistors will equal the voltage gain across the power supply (24.0 V). As the resistors have equal resistance, there will be the same voltage drop across each.

Therefore, the voltage drop across each resistor will be $\frac{24.0}{N}$ volts.

For one of the resistors (when the current is 4.00 A):

$$V = IR$$

$$\frac{24.0}{N} = 4.00 \times 2.00.$$

Therefore, $N = 3$.

SAMPLE PROBLEM 7.12

Number of resistors in a parallel circuit

Calculate how many 20.0 Ω resistors, connected in parallel to a 2.00 V power supply, would result in a current of 1.00 A through the power supply.

SOLUTION

As the resistors are in parallel with the power supply, the voltage drop across each resistor will be 2.00 V.

For one of the resistors:

$$V = IR$$

$$2.00 = I \times 20.0.$$

Therefore, $I = 0.100$ A.

The sum of the currents through the resistors will equal the current through the power supply (1.00 A). If there are N resistors, each with a current of 0.100 A, the current through the power supply will be $N \times 0.100$ A.

As the current through the power supply is 1.00 A, $N \times 0.100 = 1.00$.

Therefore, $N = 10$.

PRACTICAL ACTIVITIES

7.2

Addition of resistances

Addition of resistances

Problems involving resistors in series and parallel can always be solved by the methods used in the previous sections. However, it is sometimes easier to solve these problems by theoretically replacing all the resistors in a circuit by a single equivalent resistor. This single resistor would cause the same current to flow through the power supply as the original combination of resistors. Finding the resistance of the equivalent resistor is called addition of resistances. The equivalent resistance is also referred to as the total resistance.

The rules for addition of resistances are given here without proof. Proofs for the rules are given later in this chapter in the Physics in focus on pages 129–130.

Addition of resistances in series

If n resistors, with resistances R_1, R_2 ... R_n, are connected in series to a power supply, the combination is equivalent to a single resistor with a resistance of R_{series}, where:

$$R_{series} = R_1 + R_2 + \ldots + R_n.$$

SAMPLE PROBLEM 7.13

Adding resistances in series

Two resistors of 12.0 Ω and 18.0 Ω are connected in series to a 60.0 V power supply. Calculate the current flowing through the circuit.

SOLUTION

The two resistors are equivalent to a single resistor whose resistance is given by:

$$\begin{aligned} R_{series} &= 12.0 + 18.0 \\ &= 30.0\ \Omega. \end{aligned}$$

$$\begin{aligned} I &= \frac{V}{R} \\ &= \frac{60.0}{30.0} \\ &= 2.00\ \text{A} \end{aligned}$$

Addition of resistances in parallel

If n resistors, with resistances R_1, R_2 . . . R_n, are connected in parallel to a power supply, the combination is equivalent to a single resistor of resistance $R_{parallel}$, where:

$$\frac{1}{R_{parallel}} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}.$$

Note that when resistors are connected in parallel, the resistance of the equivalent resistor is always less than the resistance of any of the parallel resistors. Adding resistances in parallel decreases the total resistance and increases the total current.

SAMPLE PROBLEM 7.14

Adding resistances in parallel

Two resistors of resistance 12.0 Ω and 4.00 Ω are connected in parallel to a power supply of 30.0 V. Calculate what the current will be through the power supply.

SOLUTION The combination of the two resistors in parallel is equivalent to a resistor whose resistance is given by:

$$\frac{1}{R_{parallel}} = \frac{1}{R_1} + \frac{1}{R_2}$$
$$= \frac{1}{12.0} + \frac{1}{4.00}.$$

Therefore: $R_{parallel} = 3.00\ \Omega$.

$$I = \frac{V}{R}$$
$$= \frac{30.0}{3.00}$$
$$= 10.0\ \text{A}$$

PHYSICS IN FOCUS

Derivation of rules for addition of resistances

(a) Derivation of the formula for addition of resistances in series.

Consider a circuit with n resistors in series with a power supply. The resistor with resistance R_{series} will produce the same current when connected to the power supply as the series combination of resistors. This is shown in figure 7.17.

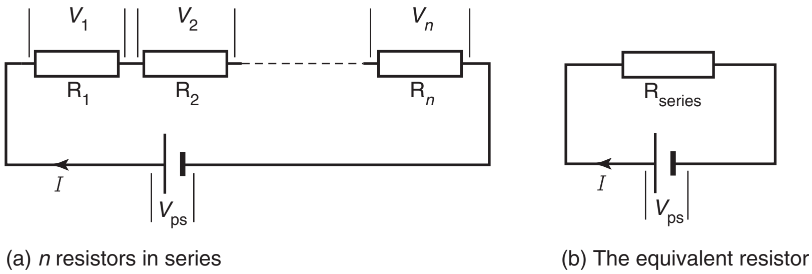


Figure 7.17 *Addition of resistances in series*

When a current, I, passes through a resistor with resistance R, the voltage drop, V, across the resistor is given by the formula: $V = IR$.

In each circuit in figure 7.17, the sum of the voltage drops across the resistors equals the voltage gain across the power supply.

In circuit (a):

$V_{ps} = IR_1 + IR_2 + \ldots + IR_n$.

In circuit (b):

$V_{ps} = IR_{series}$.

Therefore:

$IR_{series} = IR_1 + IR_2 + \ldots + IR_n$
$= I(R_1 + R_2 + \ldots + R_n)$.

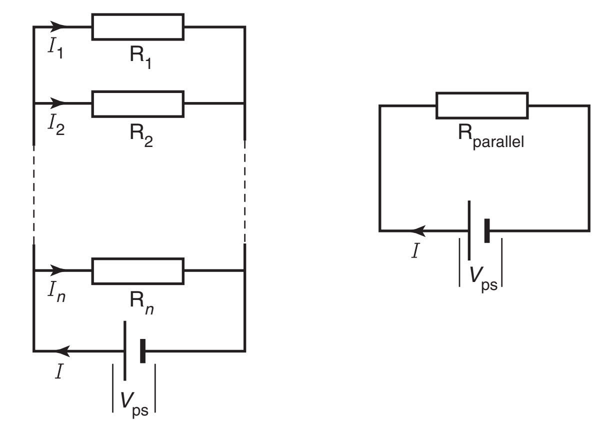
Therefore:

$R_{series} = R_1 + R_2 + \ldots + R_n$.

(continued)

(b) Derivation of the formula for the addition of resistances in parallel.

Consider a circuit with n resistors in parallel with a power supply. The resistor with resistance R_{parallel} will produce the same current when connected to the power supply as the parallel combination of resistors. This is shown in figure 7.18.



(a) n resistors in parallel (b) The equivalent resistor

Figure 7.18 *Addition of resistances in parallel*

The voltage drop across each resistor in parallel is V_{ps}. The voltage drop across R_{parallel} is also V_{ps}. When there is a voltage drop, V, across a resistor with resistance R, the current, I, through the resistor is given by the formula: $I = \dfrac{V}{R}$.

Therefore, in circuit (a):

$$I_1 = \frac{V_{ps}}{R_1}$$

$$I_2 = \frac{V_{ps}}{R_2}$$

and so on.

In circuit (b):

$$I = \frac{V_{ps}}{R_{\text{parallel}}}.$$

But, for resistors in parallel:

$$I = I_1 + I_2 + \ldots + I_n.$$

Therefore:

$$\frac{V_{ps}}{R_{\text{parallel}}} = \frac{V_{ps}}{R_1} + \frac{V_{ps}}{R_2} + \ldots + \frac{V_{ps}}{R_n}$$

$$= V_{ps}\left(\frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}\right).$$

Therefore:

$$\frac{1}{R_{\text{parallel}}} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}.$$

7.2 USING ELECTRICITY SAFELY IN THE HOME

The use of electricity in the home is potentially dangerous. The risks that must be understood and minimised are *fire* and *electric shock*.

In this section of the course we will study the supply of electricity to the home and the measures taken to ensure its safe use.

The alternating voltage supply

The neutral wire is one of the wires that delivers electricity to a household from a power station. This wire is maintained at earth potential.

The active wire is one of the wires that brings electricity to a household from a power station. This wire has a voltage that varies from +340 V to −340 V with respect to the neutral wire.

The electricity supply is delivered to a home through overhead or underground cables. The cables are connected to a switchboard where there is a switch that turns the home electricity supply on and off, and a meter that measures the amount of electrical energy used. From the switchboard, the electricity passes to a fuse box (or circuit breaker box) where it is divided into a number of parallel circuits through the home.

An alternating voltage is delivered to the home by two wires called the **neutral wire** and the **active wire** (or hot wire). The neutral wire is maintained at earth potential by being connected to the earth at the power station. (This means that there is no voltage between this wire and the earth.) The voltage of the active wire varies between +340 V and −340 V

The **peak voltage** is the maximum voltage between the active wire and the neutral wire.

with respect to the neutral wire. The maximum of 340 V is called the **peak voltage**. Figure 7.19 shows how the voltage of the active wire varies with time.

The graph of voltage against time is a sine curve.

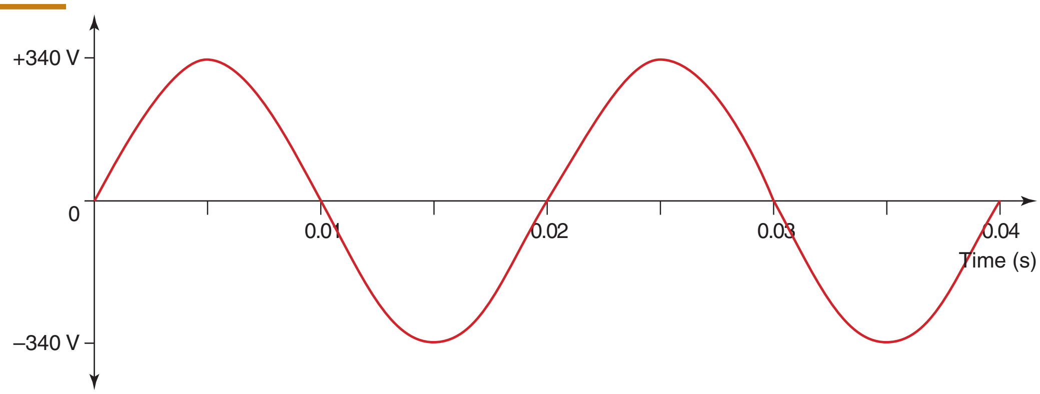


Figure 7.19 *Alternating voltage supply*

As the voltage between the active lead and the neutral lead is varying, an 'average value' is used. A DC voltage of 240 V has the same heating effect when applied across a resistor as the household AC supply has. The household supply is therefore called a 240 V AC supply, and this value can be used in all current and energy calculations involving the alternating household supply.

A **cycle** is a complete variation in the voltage.

A complete variation in the voltage is called a **cycle**. For the electricity supply a cycle takes 0.02 s. There are therefore 50 cycles per second (also referred to as 50 hertz or 50 Hz).

An alternating voltage produces an alternating current.

SAMPLE PROBLEM 7.15

Calculating current for a household electricity supply

A $2.0 \times 10^3\ \Omega$ resistor is connected between the terminals of a 240 V AC supply. What AC current will flow through the resistor?

SOLUTION

$$I = \frac{V}{R}$$
$$= \frac{240}{2.0 \times 10^3}$$
$$= 0.12 \text{ A AC}$$

Electric shock

An **electric shock** is a violent disruption of the nervous and muscular systems caused by the passage of an electric current through the body.

Electric shock, or electrocution, is caused when an electric current passes through the body and causes a violent disturbance of the nervous system. The disturbance of the nervous system produces effects on the muscles of the body.

In household situations, electric shock most commonly happens when a person who is in contact with the ground touches something that is 'live'; that is, in contact with the active wire. For example, many people have been electrocuted while cleaning the gutters on their roofs when they have accidentally touched the active wire bringing electricity to the house, thus making a connection through the ladder to the ground.

Some appliances also produce DC voltages that can cause electric shock. For example, television sets produce DC voltages of the order of 1000 V.

An electric current through the body can cause muscles to contract. If the muscles required for breathing are involved, this can prevent breathing and may cause death from lack of oxygen to the brain. The effect of mains alternating currents on the heart muscles is more complicated.

Table 7.1 gives an indication of the effects of alternating currents with a frequency of 50 Hz passing from hand to hand.

Table 7.1

CURRENT (mA)	PHYSIOLOGICAL EFFECTS
0–1	Not perceptible.
1–15	Pain at the points of entry and exit of the current. Involuntary contraction of muscles associated with the path of the current.
15–100	Inability to release grip. Muscular contractions of the chest and diaphragm. Contractions may prevent breathing and, if prolonged, may lead to death.
100–200	Fibrillation leading to death.
Over 200	Reversible stopping of the heart. Loss of consciousness.

Fibrillation is a condition in which the heart stops beating regularly and oscillates rapidly.

An electric shock can also cause burns to the body.

Fibrillation of the heart is a condition in which the heart stops beating regularly and oscillates rapidly. The heart is no longer able to pump blood, and death follows in a matter of minutes from lack of oxygen to the brain. It is thought that currents above 200 mA clamp the heart muscles and prevent fibrillation. Direct currents are less dangerous than alternating currents, as they do not lead to fibrillation.

The degree to which the heart and chest muscles are affected depends on the path of the current through the body. A current that passes through the chest affects the nerves involved in breathing and the nerves that cause the heart to beat. A shock is most dangerous if the current passes from a hand to the opposite foot. A current passing from hand to hand is less dangerous but can still be fatal. A current passing from foot to foot is usually not fatal.

How does voltage influence the severity of an electric shock?

As seen above, it is the current that passes through the body that is important in determining the severity of an electric shock. The current will depend on the voltage across the body and on the resistance of the body $\left(I = \frac{V}{R}\right)$. The resistance of the body varies greatly, and a low voltage can sometimes produce a large enough current to cause a serious shock. Voltages as low as 40 V have caused fatal shocks.

The resistance of the body is mainly due to the skin. Dry skin has a resistance of the order of 100 000 Ω, while wet skin has a much lower resistance of the order of 1000 Ω. The resistance also depends on the thickness of the skin. The internal resistance of a human body varies from 100 to 500 Ω. The skin is rarely completely dry because of perspiration and the typical resistance of the body is between 1000 and 2000 Ω.

A person receiving a shock in a home usually receives the mains voltage (240 V) across his or her body. If the resistance of the person is 2000 Ω, the current through the body will be:

$$I = \frac{240}{2000}$$
$$= 0.12 \text{ A}$$
$$= 120 \text{ mA.}$$

This can cause death by fibrillation.

Electrical fires

An **overloaded circuit** carries a current higher than the maximum safe value for which the circuit was designed.

A **short circuit** is where an active wire comes in contact with the neutral wire or is earthed.

A household circuit can cause fire if it is **overloaded** or if a **short circuit** occurs.

Overloading the electrical wiring

When a current passes through a conductor, heat is generated. The amount of heat generated increases as the current through the conductor increases. Each piece of household wiring is designed to carry a certain maximum current without overheating.

Wiring that has to carry a high current is made thicker to decrease the resistance and thus decrease the heat generated. If a wire is overloaded by passing excessive current through it, it can become sufficiently hot to cause a fire.

The diameter of the wiring and the maximum currents for each type of household circuit are shown in table 7.2.

Table 7.2

TYPE OF CIRCUIT	DIAMETER OF WIRING (mm)	MAXIMUM CURRENT (A)
Lights	2.5	8
Power points	1.0	20
Stove	6.0	32

eBook*plus*

Weblink:
Electric circuits
Follow the Electricity and Magnetism link to Electric Circuits.

Fires can also be caused by overloading extension cords. Cheaper cords usually have a lower current rating than more expensive cords.

Short circuits

A short circuit occurs when the active wire comes in contact with the neutral wire or with a conductor connected to the earth. When this happens, a circuit is formed with very little resistance, leading to a very high current. Sufficient heat can be generated to cause a fire.

Safety devices

A number of safety measures are used in household circuits. These include:

- double insulation
- fuses
- circuit breakers
- earth wires
- residual current devices.

Double insulation

All household electrical wiring must be covered with a flexible insulator such as polyvinylchloride (PVC). Where possible, electrical appliances have a casing of rigid plastic to act as insulation. Many appliances — for example, electric shavers — have double insulation. That is, an inner insulation and an outer insulation in case the inner insulation fails.

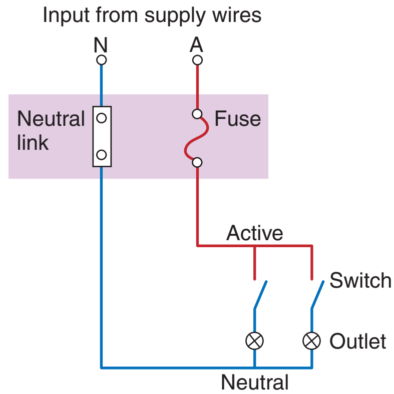


Figure 7.20 ***Connecting a fuse into a household circuit***

Fuses

Fuses are used to prevent overloading of household circuits. A fuse is made of resistance wire with a low melting point (e.g. a lead–tin alloy). The fuse is connected in series into the active wire. This is illustrated in figure 7.20.

As the fuse is in series, the total current in the circuit passes through the fuse. Because the fuse is made of higher resistance material than the wiring in the circuit, it will generate more heat and its temperature will rise higher than the temperature of the wiring. The fuse is designed to melt when the current through the circuit exceeds the limit for the wiring. (To prevent the fuse itself being a fire hazard, the fuse wire is inside an insulating, high-melting-point casing.)

Fuse wires are made with different ratings for use in different circuits. Common ratings for fuses are 8 A, 20 A and 32 A.

Because a fuse melts when the maximum current is exceeded, it is necessary to replace the fuse wire after the fuse 'blows'. It is important to use the correct rating fuse wire for each circuit. If a 20 A fuse wire is used in a circuit made of wiring with a 6 A rating, the fuse would not melt if the current exceeded 6 A, and a fire could occur.

PRACTICAL ACTIVITIES

7.3

Fuses

A diagram of a circuit breaker can be found in chapter 8 on page 160.

Circuit breakers

Circuit breakers have the same function as fuses. A circuit breaker uses an electromagnet to mechanically break the circuit when the current exceeds the maximum value. Circuit breakers have an advantage over fuses because they can be reset after they have been 'tripped'.

Earth wires

The function of the earth wire is to provide protection from electric shock for people using electric appliances.

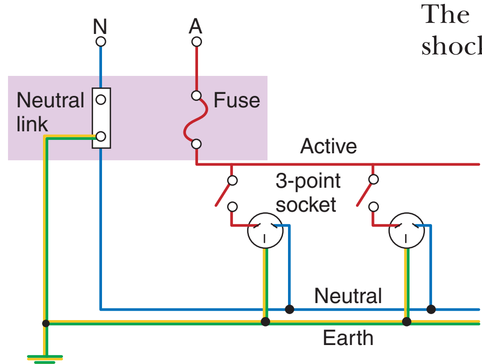


Figure 7.21 ***The earth wire and connections to a power outlet***

When household wiring is installed, a good conductor is used to connect the fuse box to the earth. At the fuse box, the wire coming from the earth is (a) connected to the neutral wire from the electricity supply and (b) connected to earth wires that go from the fuse box to each power outlet.

Power outlets have three wires connected to them from the fuse box — the active wire, the neutral wire and the earth wire. Figure 7.21 shows the earth-wire connections at the fuse box and the connections to a power outlet.

Most appliances are connected to the power outlet by a three-point plug, connected to a lead containing three wires which continue the active, neutral and earth wires to the appliance. Figure 7.22 shows a power outlet and a three-point plug.

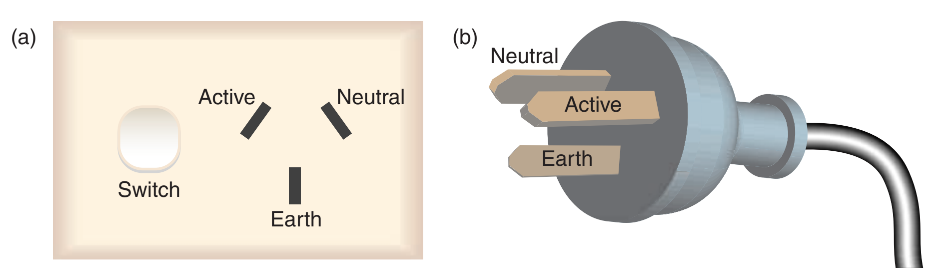


Figure 7.22 ***(a) A power outlet and (b) a three-point plug***

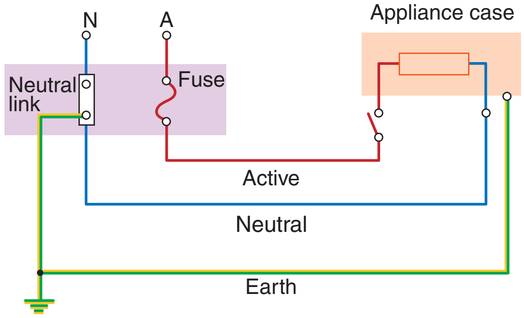


Figure 7.23 *Connections to an appliance*

The active and neutral wires provide power to the appliance. The earth wire is connected to the casing of the appliance. Figure 7.23 shows the connections of the active, neutral and earth wires to an appliance.

The earth wire provides protection in case the active wire comes in contact with the casing of the appliance. Without the earth wire, a person touching the casing could receive an electric shock, as current flowed through his or her body to the earth. With the earth wire in place, as soon as the active wire comes in contact with the casing, a large current will flow to the earth through the earth wire. This large current would cause the fuse to blow or the circuit breaker to trip, thus disconnecting the active wire from the appliance.

The fuse (or circuit breaker) is always in the active wire, so that the large current will pass through it. If the fuse (or circuit breaker) were in the neutral wire, the large current that flows to earth would not pass through it.

Appliances with double insulation do not require an earth wire. They are connected to the power outlet by leads containing only two wires, the active wire and the neutral wire.

In order to distinguish the active, neutral and earth wiring, the insulating coverings are coloured differently. The active wire is coloured brown, the neutral wire is coloured blue and the earth wire is coloured with green and yellow stripes. This is illustrated in figure 7.24.

The colours used in diagrams in this book for active, neutral and earth wires are not the colours used in actual household wiring.

Figure 7.24 *Colours of active, neutral and earth wires*

(Previously the active wire was coloured red, the neutral wire was coloured black and the earth wire was coloured green. The colours were changed to avoid mistakes due to red–green colour blindness. Some old houses have these colours in their wiring.)

Residual current devices

A residual current device (RCD) detects any leakage of current to the earth either through your body or through some other conductor. It is designed to switch off the current very quickly before it reaches a harmful level. Sometimes an RCD is called a safety switch. An RCD reduces the risk of electric shock; however, it will not protect you in all instances.

An RCD must be installed as standard fuse-box equipment in all new houses and buildings.

7.4

Building a 12 V model of a household wiring system

7.5

Designing a household wiring system

Household wiring

A typical household wiring system is shown in figure 7.25. The following points should be noted:

- there are a number of circuits, each designed to carry a certain maximum current
- each circuit has a fuse in the active wire
- there are separate circuits for light outlets and power outlets
- the stove and the hot water system each has its own circuit (not shown in diagram)
- the outlets in each circuit are connected in parallel; this means that each appliance will have 240 V across it
- there is a switch in series with each outlet; this enables each appliance to be switched on and off independently of the other appliances
- the switch is on the active side of an appliance. (If the switch were on the neutral side, the appliance could cause an electric shock even if it were switched off.)

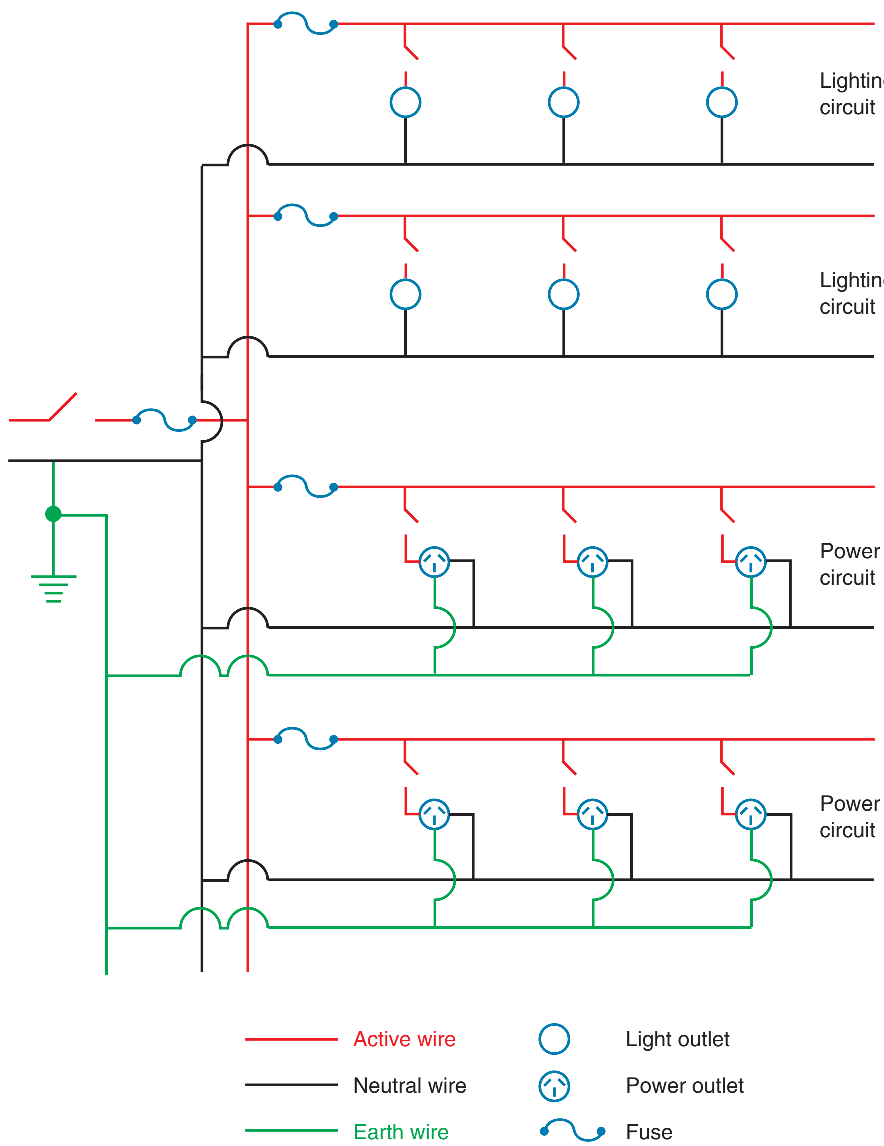


Figure 7.25 *A household wiring system*

PHYSICS IN FOCUS

RMS values of voltage and current

Consider an appliance connected to the alternating mains power supply. The voltage of the active wire changes with respect to the neutral wire. Figure 7.26 shows how the mains voltage varies with time.

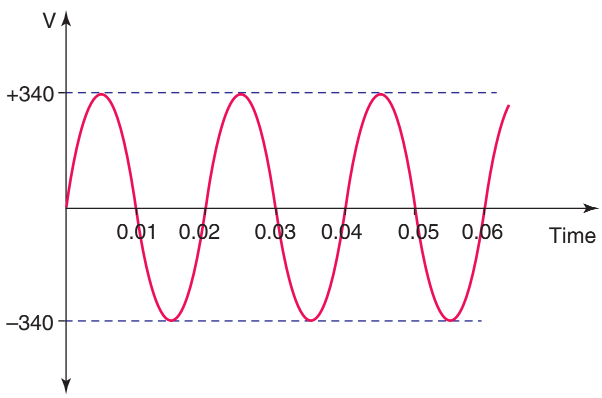


Figure 7.26 ***Graph of voltage against time***

When the voltage is positive, the (conventional) current flows from the active wire to the neutral wire. When the voltage is negative, the current flows from the neutral wire to the active wire. As the voltage varies equally between positive and negative its average value is zero. If an individual current carrier (free electron in a metal) could be observed, it would vibrate back and forth but would not progress in either direction.

To obtain a value of the voltage that can be used in calculations, the voltage is first squared. This makes all the values positive. Figure 7.27 shows the result of squaring the voltage.

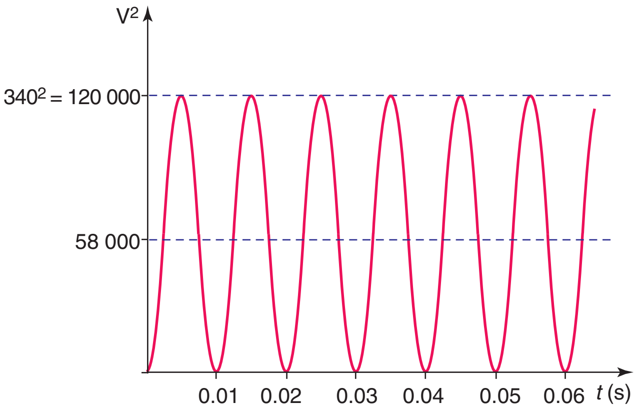


Figure 7.27 ***Graph of voltage^2 against time***

An average can now be found of the squared values. This is called the mean-square-voltage (mean is another word for average). Its value is approximately 58 000 V^2 and is marked on the vertical axis in figure 7.27.

The mean-square-voltage has the dimension *volt*^{2}. To obtain a value with dimension *volt*, the square root of the mean-square-voltage is found. This yields a value known as the root-mean-square voltage. This is usually abbreviated to RMS voltage. For the household circuit the RMS voltage is 240 V.

As the current in an AC circuit is also alternating, there will also be an RMS current value. When $I = \frac{V}{R}$ is used in an AC circuit, the RMS voltage and RMS current are used. The RMS values are also used in energy and power calculations.

SUMMARY

- When resistors are connected in series to a power supply, the same current passes through each resistor and through the power supply.
- When resistors are connected in series to a power supply, the sum of the voltage drops across the resistors equals the voltage rise across the power supply.
- When resistors are connected in parallel to a power supply, the sum of the currents through the resistors equals the current through the power supply.
- When resistors are connected in parallel to a power supply, the voltage drop across each resistor equals the voltage rise across the power supply.
- The household electricity supply is 240 V AC at a frequency of 50 Hz.
- Electricity is delivered to the household by two wires: the active or hot wire and the neutral wire.
- Household wiring is divided into a number of circuits for lighting, a number of circuits for power and special circuits, such as those for the hot water system and the stove.
- Each circuit is designed to carry a maximum current appropriate to its use.
- In each lighting and power circuit, the outlets are connected in parallel and there is a switch for each outlet.
- Each circuit has a fuse or circuit breaker of the appropriate rating inserted into the active wire.
- Hazards arising from the use of electric power are fire and electric shock.
- Fire can be caused by overloading a circuit or by a short circuit.
- Electric shock is usually caused by current from an active wire passing through the body to the earth.
- The severity of an electric shock depends on a number of factors. A shock from the mains power supply can be lethal.
- Electric shocks can be caused by the 240 V AC mains supply and by high DC voltages produced in some appliances.
- Safety devices include double insulation, fuses, circuit breakers, residual current devices and earthing.

QUESTIONS

1. Draw circuit diagrams showing resistors of 2.00 Ω, 4.00 Ω and 6.00 Ω (a) in series and (b) in parallel with a 12.0 V battery.
2. In the circuit shown in figure 7.28, the current through the 4.00 Ω resistor is 5.00 A. Calculate:
   (a) the current through the 8.00 Ω resistor
   (b) the current through the power supply.

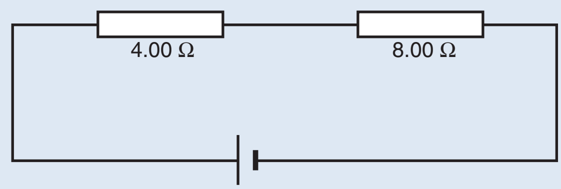


Figure 7.28

3. In the circuit shown in figure 7.29, the current through the power supply is 6.00 A and the current through R_1 is 4.00 A.
   (a) Calculate the current through R_2.
   (b) Deduce which resistor has the greater resistance.

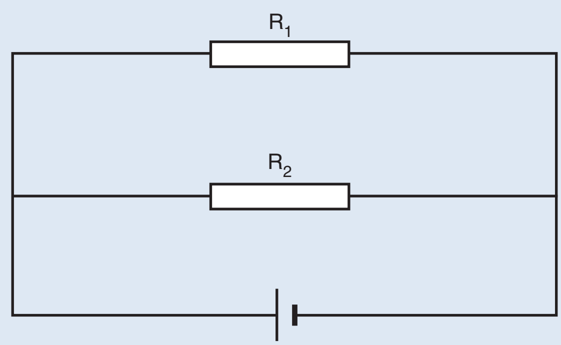


Figure 7.29

4. In the circuit shown in figure 7.30, the current through the power supply is 20.00 A and the current through the 6.00 Ω resistor is 4.00 A. Calculate the current through each of the 3.00 Ω resistors.

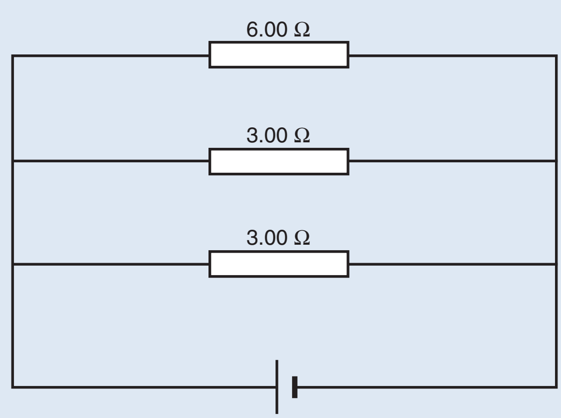


Figure 7.30

5. In the circuit shown in figure 7.31, the power supply has a voltage of 7.00 V and the voltage drop across R_1 is 4.00 V.
   (a) Calculate the voltage drop across R_2.
   (b) Deduce which resistor has the greater resistance.

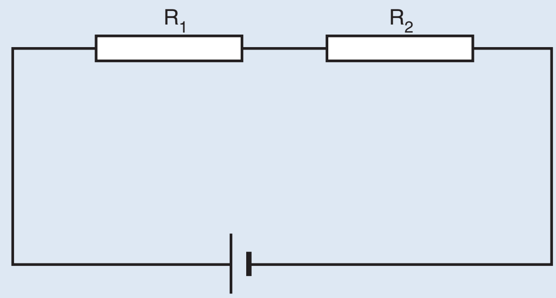


Figure 7.31

6. In the circuit shown in figure 7.32, the power supply has a voltage of 20.0 V and the voltage drop across the 6.00 Ω resistor is 12.0 V. Calculate the voltage drop across each of the 2.00 Ω resistors.

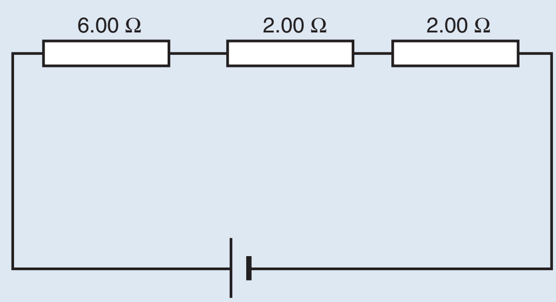


Figure 7.32

7. In the circuit shown in figure 7.33, the voltage drop across the 5.00 Ω resistor is 20.0 V. Calculate:
   (a) the voltage drop across the 8.00 Ω resistor
   (b) the voltage of the battery.

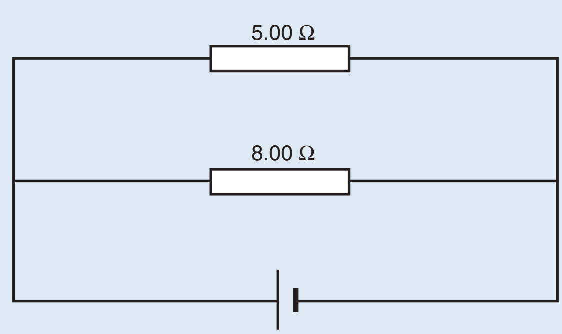


Figure 7.33

8. In the circuit shown in figure 7.34, the voltage drop across the 12.0 Ω resistor is 36.0 V. Calculate:
   (a) the current through the 12.0 Ω resistor
   (b) the current through the 4.00 Ω resistor
   (c) the voltage drop across the 4.00 Ω resistor
   (d) the voltage of the power supply.

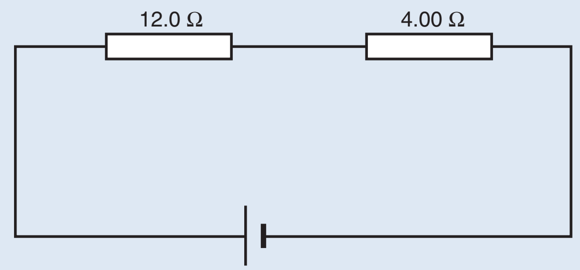


Figure 7.34

9. In the circuit shown in figure 7.35, the current through the 5.00 Ω resistor is 4.00 A. Calculate:
   (a) the voltage drop across the 5.00 Ω resistor
   (b) the voltage drop across the resistor R
   (c) the current through the resistor R
   (d) the resistance of the resistor R.

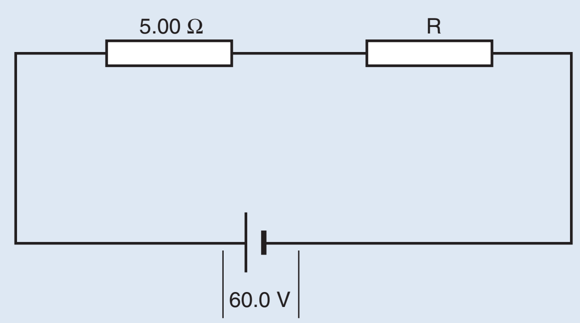


Figure 7.35

10. For the circuit shown in figure 7.36, calculate the current through the circuit.

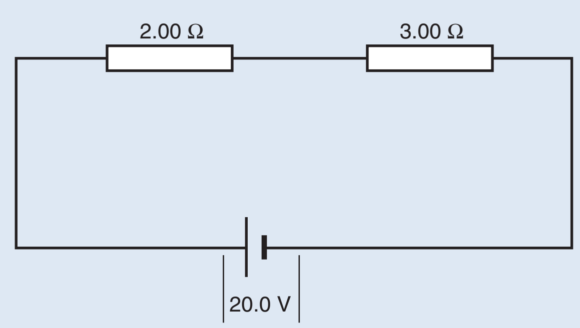


Figure 7.36

11. For the circuit shown in figure 7.37, calculate:
    (a) the potential drop across the 12.0 Ω resistor
    (b) the potential drop across the 4.00 Ω resistor
    (c) the current through the 4.00 Ω resistor
    (d) the potential gain across the power supply
    (e) the current through the power supply.

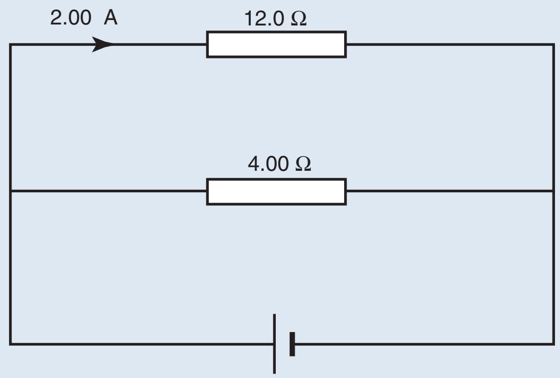


Figure 7.37

12. For the circuit shown in figure 7.38, calculate:
    (a) the current through the 25.0 Ω resistor
    (b) the current through the 15.0 Ω resistor
    (c) the current through the power supply.

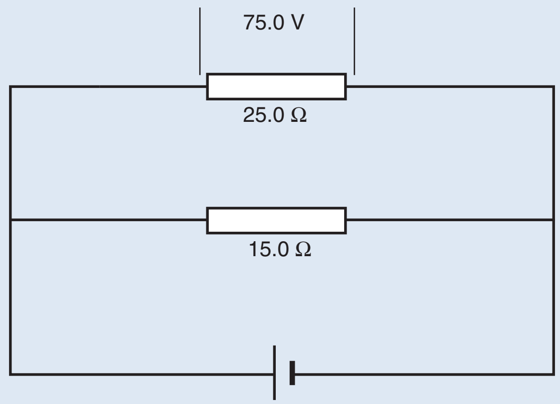


Figure 7.38

13. For the circuit shown in figure 7.39, calculate:
    (a) the total resistance
    (b) the current through the power supply
    (c) the potential drop across the 12.0 Ω resistor.

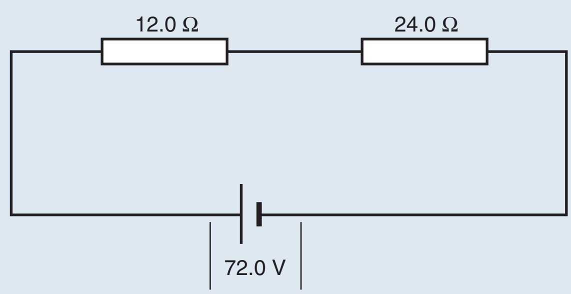


Figure 7.39

14. When a certain resistor is connected to a 25.0 V power supply, the current through the resistor is 12.5 A. Calculate the resistance of a second resistor, connected in series, that reduces the current to 5.00 A.

15. In each of the following cases, calculate the total resistance:
    (a) resistors of 12.0 Ω and 16.0 Ω in series
    (b) resistors of 12.0 Ω and 4.00 Ω in parallel
    (c) resistors of 12.0 Ω, 10.0 Ω and 17.0 Ω in series
    (d) resistors of 12.0 Ω, 10.0 Ω and 17.0 Ω in parallel
    (e) six 15.0 Ω resistors in series
    (f) five 2.00 Ω resistors in parallel.

16. In each of the following cases, calculate:
    (i) the current through each resistor
    (ii) the voltage drop across each resistor.
    (a) 5.00 Ω and 7.00 Ω resistors connected in series to a 6.00 V power supply
    (b) 10.0 Ω, 20.0 Ω and 40.0 Ω resistors connected in series to a 140.0 V power supply
    (c) 2.4 Ω and 3.7 Ω resistors connected in series to a 15 V supply
    (d) 11.2 Ω, 20.4 Ω and 31.5 Ω resistors connected in series to a 128 V power supply
    (e) 2.00 Ω and 3.00 Ω resistors connected in parallel to a 12.0 V power supply
    (f) 12.0 Ω, 24.0 Ω and 60.0 Ω resistors connected in parallel to a 48.0 V power supply
    (g) 17.3 Ω and 25.6 Ω resistors connected in parallel to a 125 V power supply
    (h) 2.53 Ω, 7.12 Ω and 4.28 Ω resistors connected in parallel to a 30.5 V power supply.

17. Explain why a fuse or a circuit breaker is connected into each circuit of a household electrical system.

18. Explain why a fuse or circuit breaker is connected into the active wire rather than into the neutral wire.

19. Explain the advantage of a circuit breaker over a fuse.

20. Explain the function of the earth wire in a household electrical system.

21. Explain the function of double insulation.

22. Discuss the use of a residual current device.

23. Tom claims that the effect of an electric shock depends on the voltage. Bill claims that the effect depends on the current. Explain the confusion.

24. Explain why a more serious shock is experienced if the skin is wet than if it is dry.

7.1(a) CURRENT IN A SERIES CIRCUIT

Aim

To investigate current in a series circuit

Apparatus

power supply
ammeter
two resistors, R_1 and R_2, with different resistances
switch
connecting wires

Theory

In a series circuit the same current flows through each resistor.

Method

1. Connect the apparatus as shown in figure 7.40.
2. Connect the ammeter in turn to measure the current at each of the points X, Y and Z, when the switch is closed.

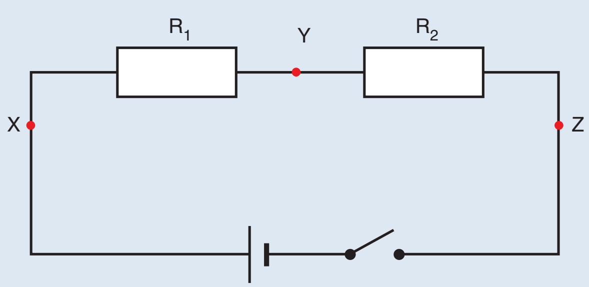


Figure 7.40

Results

Record the current measurements as in the table below.

Current at X	
Current at Y	
Current at Z	

Analysis

1. What is the current passing through the power supply?
2. What is the current passing through R_1?
3. What is the current passing through R_2?
4. Write a sentence to sum up what you have observed about the current in a series circuit.

Questions

1. Does current decrease as it passes through a resistor?
2. What does decrease as a current passes through a resistor?
3. Does current increase as it passes through a power supply?
4. What does increase as a current passes through a power supply?

7.1(b) CURRENT IN A PARALLEL CIRCUIT

Aim

To investigate current in a parallel circuit

Apparatus

power supply
ammeter
two resistors, R_1 and R_2, with different resistances
switch
connecting wires

Theory

When current passes through resistors in parallel, the current passes partly through each of the resistors. The current through the power supply is equal to the sum of the currents through the resistors.

Method

1. Connect the apparatus as shown in figure 7.41.

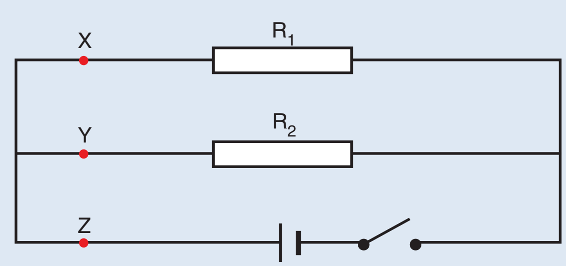


Figure 7.41

2. Connect an ammeter in turn to measure the current at each of the points X, Y, and Z, when the switch is closed.

Results

Record your results as in the table on page 142.

Current at X	
Current at Y	
Current at Z	

Analysis

1. What is the sum of the currents through the two resistors in parallel?
2. Compare this with the current through the power supply.
3. Write a sentence to sum up what you have observed about currents in a parallel circuit.

Questions

1. Which resistor had the higher current passing through it?
2. What is the ratio of current through R_1 to the current through R_2?
3. What is the ratio of R_1 to R_2?
4. Comment on your answers to questions 2 and 3.

7.1(c) VOLTAGE IN A SERIES CIRCUIT

Aim

To investigate voltage in a series circuit

Apparatus

power supply
voltmeter
two resistors, R_1 and R_2, with different resistances
switch
connecting wires

Theory

In a series circuit, the voltage rise across the power supply equals the sum of the voltage drops across the resistors.

Method

1. Connect the circuit as in figure 7.42.
2. Connect the voltmeter in turn across the power supply and each of the resistors and record the readings when the switch is closed.

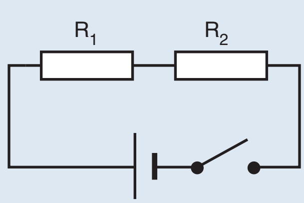


Figure 7.42

Results

Record the measurements as in the table below.

V across power supply	
V across R_1	
V across R_2	

Analysis

1. What is the sum of the voltage drops across the resistors?
2. Compare this with the voltage gain across the power supply.
3. Write a sentence to state the relationship between the voltage across the power supply and the voltage drops across the resistors.

Questions

1. Which resistor had the higher voltage drop across it?
2. What is the ratio of the voltage drop across R_1 to the voltage drop across R_2?
3. What is the ratio of R_1 to R_2?
4. Compare your answers to questions 2 and 3 and make a comment.

7.1(d) VOLTAGE IN A PARALLEL CIRCUIT

Aim

To investigate the voltage in a parallel circuit

Apparatus

power supply
voltmeter
two resistors, R_1 and R_2, with different resistances
switch
connecting wires

Theory

In a parallel circuit, the voltage drop across each resistor is equal to the voltage gain across the power supply.

Method

1. Connect the apparatus according to the circuit diagram in figure 7.43.

2. Connect the voltmeter in turn to measure the voltage (with the switch closed) across the power supply and across each resistor.

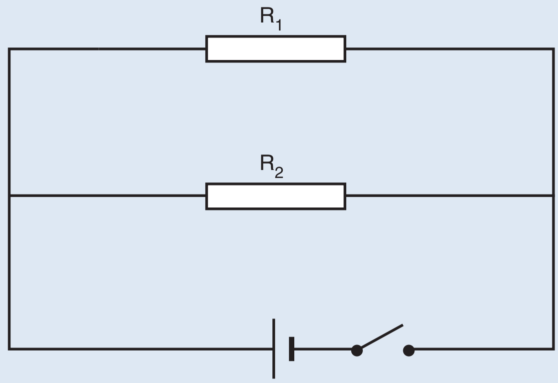


Figure 7.43

Results

Record your results as in the table below.

V across power supply	
V across R_1	
V across R_2	

Analysis

Write a sentence to sum up these results.

Questions

1. Does any current pass through *both* R_1 and R_2?
2. Consider *only* that part of the current that passes through R_1.
   (a) What is its voltage gain in the power supply?
   (b) What is its voltage loss in R_1?
3. Consider *only* that part of the current that passes through R_2.
   (a) What is its voltage gain in the power supply?
   (b) What is its voltage loss in R_2?
4. Consider the *total* current in the circuit.
   (a) What is its voltage gain in the power supply?
   (b) What is its voltage loss in the resistors?

7.2 ADDITION OF RESISTANCES

Aim

To use the laws of addition of resistances to predict the resistance of two resistors connected (a) in series and (b) in parallel and to test the predictions experimentally

Apparatus

power supply
variable resistor
ammeter
voltmeter
two resistors, R_1 and R_2, with different resistances
switch
connecting wires

Theory

The total resistance, R_{series}, of two resistors, R_1 and R_2, connected in series, is given by the rule:

$$R_{series} = R_1 + R_2.$$

The total resistance, $R_{parallel}$, of two resistors, R_1 and R_2, connected in parallel, is given by the rule:

$$\frac{1}{R_{parallel}} = \frac{1}{R_1} + \frac{1}{R_2}.$$

To measure the total resistance of a combination of resistors, measure the voltage drop, *V*, across the combination, and the current, *I*, passing through the combination. The total resistance is given by the formula: $R = \dfrac{V}{I}$.

Method

1. Connect the circuit as shown in figure 7.44. Between the points X and Y, connect in turn:
   (a) R_1
   (b) R_2
   (c) R_1 and R_2 in series
   (d) R_1 and R_2 in parallel.

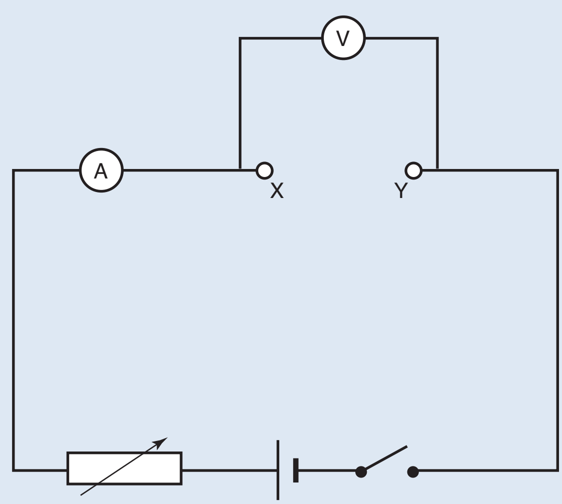


Figure 7.44

2. In each case adjust the voltage of the power supply and the resistance of the variable resistor so that the current has the value suggested by your teacher.
3. With the switch closed, measure the total current passing from X to Y and the voltage across XY.

Results

Record your results as in the table below.

	V (V)	I (A)	$R = \frac{V}{I}$ (Ω)
R_1			
R_2			
R_1 and R_2 in series			
R_1 and R_2 in parallel			

The third column is calculated from the measurements of voltage and current. This column gives the *experimental* values of the resistances and their series and parallel combinations.

Analysis

1. Using the laws of addition of resistances, calculate the *theoretical* values R_{series} and $R_{parallel}$. (Use the values of R_1 and R_2 found experimentally, not the nominal values written on the resistors.)
2. Compare the theoretical and experimental values by completing the following table.

	EXPERIMENTAL VALUE OF RESISTANCE (Ω)	THEORETICAL VALUE OF RESISTANCE (Ω)
R_1 and R_2 in series		
R_1 and R_2 in parallel		

3. Comment on any differences between the experimental and theoretical values.

7.3 FUSES

Aim

To make a model fuse which will pass a current of 2 A but not a current of 5 A

Theory

For wires of the same material and of the same length, the thinner the wire the greater its resistance and the greater the rate at which heat is generated in it.

Apparatus

power supply
variable resistor
switch
alligator clips
steel wool
ammeter
connecting wires

Method

1. Connect the apparatus as shown in figure 7.45.

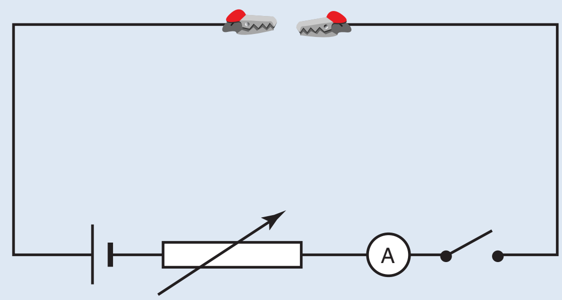


Figure 7.45

2. Take one strand, about 5 cm long, of the steel wool and insert it in the circuit as a fuse between the alligator clips. Suspend your fuse so that it doesn't make contact with any other object.
3. Start with no current through the circuit and gradually increase the current until the fuse blows. Note the current at which the fuse blows.
4. Increase the thickness of the fuse by twisting more than one strand together. (Keep the length the same.)
5. Continue using increasing numbers of strands until you have a fuse which will reliably pass a current of 2 A but will always blow if a current of 5 A is passed through it.

Results

Record your results as in the table below.

NUMBER OF STRANDS OF STEEL WOOL	MAXIMUM CURRENT WITHOUT FUSING (A)

Question

On the basis of your experiment which would you expect to be thicker, a 2 A or a 10 A fuse wire?

7.4 BUILDING A 12 V MODEL OF A HOUSEHOLD WIRING SYSTEM

Aim

To build a model of household wiring system

Apparatus

Minimum apparatus required:
- power supply
- three 12 V light globes with holders
- three switches
- connecting wires

Method

The model can be as elaborate as you wish but should contain as a minimum a circuit with three lights.

In designing your circuit, refer to the section on household wiring on page 136, and in particular to figure 7.25.

Analysis

Draw a circuit diagram of your model.

Question

What are some ways in which your model differs from an actual wiring system?

Warning

Do not attempt to build a model using 240 volts. Do not attempt any household wiring yourself.

7.5 DESIGNING A HOUSEHOLD WIRING SYSTEM

Aim

To design the wiring system for a flat consisting of a living room, kitchen, bedroom and bathroom

Method

1. Work out how many outlets for lights and how many outlets for power points you will need. The stove and the hot water heater will each have a separate circuit.
   How many light switches will you need?
   How many lighting circuits will you need?
   How many three-point power outlets will you need?
   How many power circuits will you need?
2. The table below shows the current required for various appliances.

APPLIANCE	APPROX. CURRENT (A)
Appliance using small electric motor	0.75
Air conditioner	3–10
Clothes dryer	10
Vacuum cleaner	1–5
Hot water service	10–15
Iron	4–5
Kettle or jug	7–13
Light (100 W)	0.5
Radio	0.25
Radiator	4–10
Refrigerator	1–2
Television set	1
Toaster	3–5
Washing machine	2–12

   Refer to the table to determine the rating of the fuse for each circuit (8A, 20A, 32A).
3. Draw a plan of the unit showing the wiring. (Put the fuse box in the kitchen.)

CHAPTER 8 USING ELECTRICITY IN THE HOME

Figure 8.1 *Some of the wide variety of electrical appliances available for use in Australian households*

Remember

Before beginning this chapter, you should be able to:

- describe magnetism in terms of the ability of magnets to attract iron
- state that the magnetism of a magnet is concentrated at the ends called the poles
- state that each magnet has a north pole and a south pole and that if a magnet is freely suspended the north pole of the magnet will point north
- describe the structure of an electromagnet
- understand some of the everyday applications of magnets, electromagnets and magnetic strips.

Key content

At the end of this chapter you should be able to:

- explain that power is the rate at which energy is transformed from one form to another
- identify the relationship between power, potential difference (voltage) and current
- solve problems and analyse information using $P = VI$
- identify that the total amount of energy used depends on the length of time that the current has been flowing, and solve problems using $Energy = VIt$
- explain why the kilowatt-hour is used to measure electrical energy consumption rather than the joule
- describe the behaviour of the magnetic poles of a bar magnet when they are brought close together
- define the direction of the magnetic field at a point as the direction of the force on a very small north magnetic pole placed at that point
- describe the magnetic fields around pairs of magnetic poles
- describe the production of a magnetic field by an electric current in a straight conductor
- describe how the right-hand grip rule can be used to determine the direction of current and field lines
- compare the nature and generation of magnetic fields by bar magnets and solenoids
- describe the direction of current and magnetic field using the symbols $\otimes$ and $\odot$
- explain one application of magnetic fields in households.

In this chapter, the principles involved in the operation of many domestic appliances will be studied. Most domestic appliances depend on heating effects of currents or on electric motors.

8.1 POWER IN ELECTRIC CIRCUITS

PRACTICAL ACTIVITIES

8.1

Model heating coil

In this book the symbol 'W' is used for a quantity of energy, as in '$W = 20$ J'. The symbol 'W' is used for a unit of power, as in '$P = 100$ W'. This distinction is not made in handwriting.

Meaning of power

Power is the rate at which energy is transformed from one form to another. If energy, W, is transformed in time, t, then the power, P, is given by the formula:

$$P = \frac{W}{t}.$$

Power is the rate at which energy is transformed from one form into another.

The **watt** (W) is the SI unit of power.

The SI unit of power is the **watt** (W). One watt is the same as one joule second^{-1}. Other units of power are the kilowatt (kW) and the megawatt (MW), where 1 kW = 10^3 W and 1 MW = 10^6 W.

For example:

1. a 1000 W electric heater transforms 1000 J of electric potential energy into heat energy in 1 s
2. a battery generating 12 W converts 12 J of chemical energy into electric potential energy in 1 s.

SAMPLE PROBLEM 8.1

Calculating power (given energy and time)

Calculate the power generated by a power supply that produces 5.0×10^3 J of electric potential energy in 2.0×10^1 s.

SOLUTION

$$\begin{aligned} P &= \frac{W}{t} \\ &= \frac{5.0 \times 10^3}{2.0 \times 10^1} \\ &= 2.5 \times 10^2 \text{ W} \end{aligned}$$

SAMPLE PROBLEM 8.2

Calculating energy (given power and time)

Calculate the electric potential energy dissipated in 1.00 h by a 2.00 kW electric heater.

SOLUTION

$$\begin{aligned} P &= 2.00 \text{ kW} \\ &= 2.00 \times 10^3 \text{ W} \\ t &= 1.00 \text{ h} \\ &= 1.00 \times 3600 \text{ s} \\ P &= \frac{W}{t} \\ 2.00 \times 10^3 &= \frac{W}{1.00 \times 3600} \end{aligned}$$

Therefore, $W = 7.20 \times 10^6$ J.

SAMPLE PROBLEM 8.3

Calculating time (given power and energy)

An electric jug has a rating of 1.00×10^3 W. If it requires 4.00×10^5 J of heat energy to heat the water in the jug to boiling point, calculate the time it will take.

SOLUTION

$$\begin{aligned} P &= \frac{W}{t} \\ 1.00 \times 10^3 &= \frac{4.00 \times 10^5}{t} \end{aligned}$$

Therefore, $t = 4.00 \times 10^2$ s.

Power, current and voltage

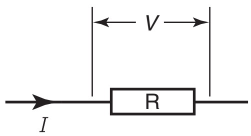


Figure 8.2 *Power dissipated in a resistor*

Consider a resistor that has a voltage drop, V, across it and a current, I, passing through it (see figure 8.2).

The current is I amps, therefore I coulombs of charge pass through the resistor in one second. Each coulomb of charge passing through the resistor dissipates V joules of electric potential energy. Therefore, in one second, VI joules of electric energy are dissipated. Power dissipated in a resistor is the number of joules of energy dissipated in one second. Therefore,

$$P = VI.$$

If the current passes for time, t, the energy, W, dissipated is:

$$\begin{aligned} W &= Pt \\ &= VIt. \end{aligned}$$

Similarly, the power, P, generated by a power supply with a voltage rise, V, across it, and current, I, passing through it, is given by the same formula:

$$P = VI.$$

The energy, W, generated in a power supply is given by the formula:

$$W = VIt.$$

SAMPLE PROBLEM 8.4

Calculating power (given voltage and current)

The voltage drop across a resistor is 2.0×10^2 V. The current passing through the resistor is 5.0 A. Calculate the power dissipated in the resistor.

SOLUTION

$$\begin{aligned} P &= VI \\ &= (2.0 \times 10^2) \times 5.0 \\ &= 1.0 \times 10^3 \text{ W} \end{aligned}$$

SAMPLE PROBLEM 8.5

Calculating current (given voltage and power)

A light globe has a power of 1.00×10^2 W and the voltage drop across it is 2.40×10^2 V. Calculate the current passing through the light globe.

SOLUTION

$$P = IV$$

$$1.00 \times 10^2 = I \times (2.40 \times 10^2)$$

Therefore, $I = 4.2 \times 10^{-1}$ A.

SAMPLE PROBLEM 8.6

Calculating energy (given current, voltage and time)

The current through a power supply is 2.0 A. The voltage rise across the power supply is 1.2×10^1 V. If the current passes for 1.0 min, calculate the electric potential energy generated by the power supply.

SOLUTION

$$\begin{aligned} W &= VIt \\ &= (1.2 \times 10^1) \times 2.0 \times (1.0 \times 60) \\ &= 1.4 \times 10^3 \text{ J} \end{aligned}$$

SAMPLE PROBLEM 8.7

Calculating voltage (given current, time and energy)

A current of 6.0 A passes through a resistor for 1.2×10^1 s and dissipates 3.20×10^2 J of electric potential energy. Calculate the voltage drop across the resistor.

SOLUTION

$$W = VIt$$

$$3.20 \times 10^2 = V \times 6.0 \times (1.2 \times 10^1)$$

Therefore, $V = 4.4$ V.

Power and resistance

When current passes though a resistor, electric potential energy is converted into heat and light energy. This is the basis of most household appliances that produce heat and light. Examples include incandescent lamps, electric jugs and toasters.

The following sample problems illustrate how to calculate power for a given resistance.

SAMPLE PROBLEM 8.8

Calculating power (given current and resistance)

A current of 2.0 A passes through a 2.5×10^{-1} Ω resistor. Calculate the power dissipated.

SOLUTION

$$\begin{aligned} V &= IR \\ &= 2.0 \times (2.5 \times 10^{-1}) \\ &= 5.0 \times 10^{-1}\,\text{V} \\ P &= VI \\ &= (5.0 \times 10^{-1}) \times 2.0 \\ &= 1.0\ \text{W} \end{aligned}$$

SAMPLE PROBLEM 8.9

Calculating power (given voltage and resistance)

There is a voltage drop of 1.00×10^{2} V across a 5.00 Ω resistor. Calculate the power dissipated.

SOLUTION

$$\begin{aligned} I &= \frac{V}{R} \\ &= \frac{1.00 \times 10^{2}}{5.00} \\ &= 0.20 \times 10^{2}\ \text{A} \\ P &= VI \\ &= (1.00 \times 10^{2}) \times (0.20 \times 10^{2}) \\ &= 2.00 \times 10^{3}\,\text{W} \end{aligned}$$

Power in series and parallel circuits

The fact that a resistor has a small resistance does not always mean that a small amount of power will be generated by a current passing through it. The following sample problem makes this clear.

SAMPLE PROBLEM 8.10

Comparing power generated in different resistances

Consider a 1.2×10^{1} V battery with (a) a 2.0 Ω resistor and (b) a 2.0×10^{-3} Ω resistor connected across the terminals. Calculate the power generated in each case.

SOLUTION

(a)
$$\begin{aligned} I &= \frac{V}{R} \\ &= \frac{1.2 \times 10^{1}}{2.0} \\ &= 6.0\ \text{A} \\ P &= VI \\ &= (1.2 \times 10^{1}) \times 6.0 \\ &= 7.2 \times 10^{1}\ \text{W} \end{aligned}$$

(b)
$$\begin{aligned} I &= \frac{V}{R} \\ &= \frac{1.2 \times 10^{1}}{2.0 \times 10^{-3}} \\ &= 6.0 \times 10^{3}\ \text{A} \\ P &= VI \\ &= (1.2 \times 10^{1}) \times (6.0 \times 10^{3}) \\ &= 7.2 \times 10^{4}\ \text{W} \end{aligned}$$

The correct relationship between resistance and power can be stated:

1. If the currents through two resistors are equal, the resistor with the greater resistance will have the greater power dissipation.
2. If the voltage drops across two resistors are equal, the resistor with the greater resistance will have the smaller power dissipation.

The first of these rules is relevant when considering resistors in series. The second rule is relevant when considering resistors in parallel (see figure 8.3).

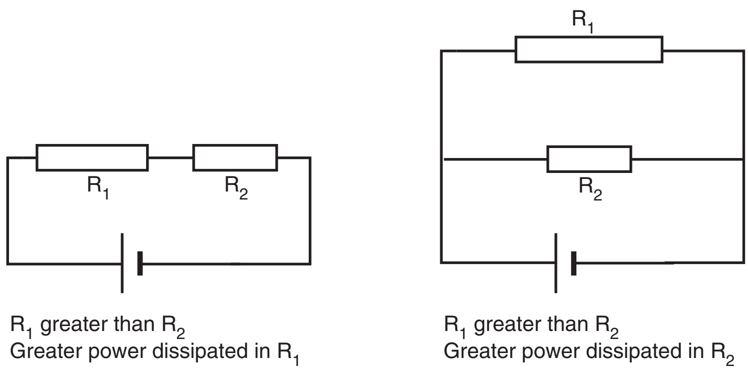


Figure 8.3 *Power in series and parallel circuits*

The following two sample problems illustrate these rules.

SAMPLE PROBLEM 8.11

Power in a series circuit

Two resistors, R_1 with resistance of 2.0 Ω and R_2 with resistance of 3.0 Ω, are connected in series to a 1.00×10^2 V power supply. Calculate the power generated by the power supply and the power dissipated in each of the resistors.

SOLUTION As the resistors are in series, the total resistance, R, is:

$$R = 2.0 + 3.0$$
$$= 5.0\ \Omega.$$

Let I be the current in the circuit:

$$V_{ps} = IR$$
$$1.00 \times 10^2 = I \times 5.0.$$

Therefore, $I = 2.0 \times 10^1$ A.

$$\text{Power generated in the power supply} = V_{ps} I$$
$$= (1.00 \times 10^2) \times (2.0 \times 10^1)$$
$$= 2.0 \times 10^3 \text{ W}$$

$$V_1 = IR_1$$
$$= (2.0 \times 10^1) \times 2.0$$
$$= 4.0 \times 10^1 \text{ V}$$

$$\text{Power dissipated in } R_1 = V_1 I$$
$$= (4.0 \times 10^1) \times (2.0 \times 10^1)$$
$$= 8.0 \times 10^2 \text{ W}$$

$$V_2 = IR_2$$
$$= (2.0 \times 10^1) \times 3.0$$
$$= 6.0 \times 10^1 \text{ V}$$

$$\text{Power dissipated in } R_2 = V_2 I$$
$$= (6.0 \times 10^1) \times (2.0 \times 10^1)$$
$$= 1.2 \times 10^3 \text{ W}$$

Note that the resistor with the greater resistance has the greater power dissipated.

Note also that the power generated in the power supply (2.0×10^3 W) is equal to the total power dissipated in the two resistors (8.0×10^2 W and 1.2×10^3 W).

SAMPLE PROBLEM 8.12

Power in a parallel circuit

Two resistors, R_1 of resistance $4.0\ \Omega$ and R_2 of resistance $3.0\ \Omega$, are connected in parallel across the terminals of a 1.2×10^1 V battery. Calculate the power generated by the battery and the power dissipated in each of the resistors.

SOLUTION

Since the resistors are connected in parallel, each resistor will have a voltage drop across it of 1.2×10^1 V.

Therefore:

$$I_1 = \frac{V}{R_1} = \frac{1.2 \times 10^1}{4.0} = 3.0\text{ A} \qquad I_2 = \frac{V}{R_2} = \frac{1.2 \times 10^1}{3.0} = 4.0\text{ A}$$

Note that the greater resistance has the smaller current passing through it.

The current, I, through the battery is given by:

$$\begin{aligned} I &= I_1 + I_2 \\ &= 3.0 + 4.0 \\ &= 7.0\text{ A.} \end{aligned}$$

Power generated by the battery is given by:

$$\begin{aligned} P &= V_{ps}I \\ &= (1.2 \times 10^1) \times 7.0 \\ &= 8.4 \times 10^1\text{ W.} \end{aligned}$$

$$\begin{aligned} \text{Power dissipated by } R_1 &= VI_1 \\ &= (1.2 \times 10^1) \times 3.0 \\ &= 3.6 \times 10^1\text{ W} \\ \text{Power dissipated by } R_2 &= VI_2 \\ &= (1.2 \times 10^1) \times 4.0 \\ &= 4.8 \times 10^1\text{ W} \end{aligned}$$

Note that the resistor with the smaller resistance generates the greater power.

Note also that the power generated by the battery (8.4×10^1 W) is equal to the sum of the powers dissipated in the resistors (3.6×10^1 W and 4.8×10^1 W).

Summary of information about resistors in series and parallel

When a number of resistors are connected in *series* to a power supply:

- the same current flows through the power supply and all the resistors
- the voltage gain across the power supply equals the sum of the voltage drops across the resistors
- the greater the resistance of a resistor, the greater the voltage drop across it
- the greater the resistance of a resistor the greater the power dissipated in it

Some home appliances use electrical energy in other ways besides converting it into heat energy. For example, electric motors, found in many appliances, convert electric potential energy into mechanical energy; radios and television sets convert electric potential energy into sound energy. The calculation of the power dissipated in these appliances is beyond the scope of this course.

- the power generated in the power supply equals the sum of the powers dissipated in the resistors.

When a number of resistors are connected in *parallel* to a power supply:

- the current through the power supply equals the sum of the currents through the resistors
- the voltage drop across each resistor is the same, and is equal to the voltage rise across the power supply
- the greater the resistance of a resistor, the smaller the current that passes through it
- the greater the resistance of a resistor, the smaller the power dissipated in it
- the power generated in the power supply equals the sum of the powers dissipated in the resistors.

Measuring domestic consumption

Electricity companies charge users for the amount of electrical energy they use. Because of the large amounts of energy involved, it is inconvenient to measure this energy in joules. For example, a 1.0 kW heater running for 2.0 hours uses 7.2×10^6 J of electrical energy.

A different unit of energy, the kilowatt-hour, is used in household electricity accounts.

A **kilowatt-hour** (kW-h) is a unit of energy equal to the amount of energy used by a 1 kW device in 1 hour.

A **kilowatt-hour** (kW-h) is defined as the amount of energy used by a 1 kW device in 1 hour. To calculate the energy used in kilowatt-hours, the following formula is used:

energy in kilowatt-hours = power in kilowatts × time in hours.

Figure 8.4 shows an electricity account for a household electricity supply.

SAMPLE PROBLEM 8.13

Calculating energy in kilowatt-hours

Calculate the energy that a 2.0 kW heater uses in 3.0 hours.

SOLUTION

$$\text{Energy in kilowatt-hours} = \text{power in kilowatts} \times \text{time in hours} = 2.0 \times 3.0 = 6.0 \text{ kW-h}$$

SAMPLE PROBLEM 8.14

Calculating energy in kilowatt-hours

Calculate the energy that a 1.0×10^2 W light uses in 8.0 hours.

SOLUTION

Convert the power to kilowatts.

$$1.0 \times 10^2 \text{ W} = \frac{1.0 \times 10^2}{1000} \text{ kW} = 1.0 \times 10^{-1} \text{ kW}$$

$$\text{Energy in kilowatt-hours} = \text{power in kilowatts} \times \text{time in hours} = (1.0 \times 10^{-1}) \times 8.0 = 8.0 \times 10^{-1} \text{ kW-h}$$

Conversions between kilowatt-hours and joules

Since the conversion factors are a matter of definition rather than of measurement, their values are known exactly and the question of significant figures does not arise.

One kilowatt-hour is the energy transformed by a one kilowatt device in one hour. That is, one kilowatt-hour is the energy transformed by a 1000 W device in 3600 seconds.

$$W = Pt = 1000 \times 3600 = 3\,600\,000 \text{ J}$$

Therefore, 1 kilowatt-hour = 3 600 000 J.

(a)

Enquiries
(Mon-Fri 8am-8pm,
Sat 8.30am-12noon) **13 15 35**

Electricity Emergencies (24 hrs)	13 13 88
Streetlight Faults	1800 044 808

www.energy.com.au

EnergyAustralia
ABN 67 505 337 385

Tax Invoice

Account Number
999 999 999

Due by
14 November 2007

Amount Payable
$410.10

BN00001 - 0001 - 07303
Mr A B Sample
1a Sample St
SAMPLE TOWN NSW 9999

Electricity Account

Location: 1a Sample St, SAMPLE TOWN NSW 9999

Total amount payable of your last bill dated 26 July 2007		**457.55**
Payments - Thankyou		-457.55 cr
Electricity (27/07/2007 to 26/10/2007)	372.82	
Subtotal of charges before GST	372.82	
Total GST payable 10%	37.28	
Total charges including GST		410.10
Total Amount Payable		**$410.10**

(b)

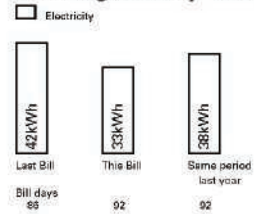
Interpreter Service 13 14 50 (Please call Mon-Fri 9am - 5pm)

للحصول على خدمات الترجمة إتصل بالرقم المدرج أعلاه.
如需傳譯員服務，請撥以上電話。
Muốn liên lạc với sở thông dịch, xin vui lòng gọi số điện thoại trên đây.
Per il servizio interpreti chiamare il numero indicato sopra.
Para comunicarse con el servicio de interpretación llame al número indicado arriba.
Για υπηρεσία διερμηνέων, τηλεφωνήστε στον παραπάνω αριθμό.

For your Information

* If you are having difficulties paying, we operate a payment plan and a customer hardship program. You may be entitled to a pensioner or life support rebate under the NSW Government funded energy rebates scheme. For details and to apply call **13 15 35.**
* If you're concerned about the planet, becoming a PureEnergy customer is an opportunity to support the future of renewable electricity generation. Please call us to become a PureEnergy customer today.

Average Daily Usage



Consider reducing your environmental impact. Greenhouse gas released to produce your energy this period ≈ 2864.4kg of CO_2

Energy Used & Costs

METER ID	THIS READING	-	LAST READING	=	ENERGY USED	X	USAGE SPLIT		X	RATE	=	COST
Off Peak 2 Hot Water - Extended Hours (27/07/07 - 26/10/07)												
HUN999999/001	37180.0		36316.0		864.0 kWh					7.9000c		$68.26
Domestic Electricity (27/07/07 - 26/10/07)												
HUN999999/001	47996.0		45816.0		2,180.0 kWh		First	1,764.0 *		11.7000c		$206.39
							Next	416.0		16.3000c		$67.81
Electricity SAC					92 Days					33.0000c/Day		$30.36
* based on 19.1781 kWh/billing day												
Total Electricity before GST					3044.0 kWh							$372.82

Figure 8.4 *(a) A section of the front part of an electricity account (b) A section of the back part of an electricity account*

PHYSICS IN FOCUS

Energy ratings for domestic appliances

Manufacturers label appliances with voltage and wattage. For example, an electric refrigerator may be labelled '240 V 200 W'. This means that the refrigerator is designed to work using a 240 V power supply and, when it is used at this voltage, it will use electrical energy at the rate of 200 joules per second.

Suppose a refrigerator is labelled 200 W. This label does not tell the purchaser how much energy the refrigerator will use, as it does not say how long the refrigerator must run each day to keep the contents cold.

In addition, this label does not tell the purchaser how efficient the appliance is. An efficient appliance is one that does not waste energy. For example, a refrigerator with poor door seals would not be very efficient, as energy would be wasted in cooling the air outside the refrigerator.

To provide more information for purchasers, Federal and State governments legislated to introduce a system of energy rating labels. Figure 8.5 shows an example of an energy rating label.

The energy rating label has two main features:

1. the comparative energy consumption. This is measured in kilowatt-hours year^{-1} and is an assessment of the annual energy consumption of the appliance based on the typical household use of the appliance.
2. the star rating. This is an assessment of the energy efficiency of the appliance.

The more stars an appliance is given, the greater its efficiency compared to similar appliances. The star rating has a minimum of one star and a maximum of six stars, shown in half-star increments.

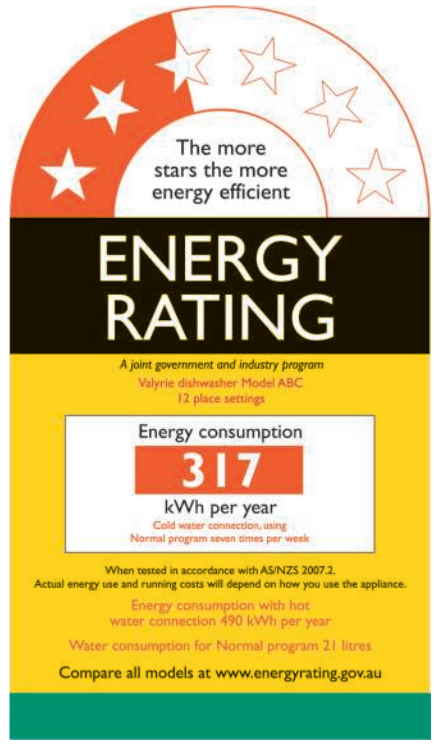


Figure 8.5 *An example of an energy rating label*

8.2 MAGNETISM

The words *magnet*, *magnetism* and *magnetic* are derived from the name of a district in Greece called Magnesia. By 600 BC, the Greeks had discovered a mineral there, now called magnetite, with the property of attracting iron.

In any sample of the mineral, the property of attracting iron is concentrated in two regions called the poles. If the sample of mineral is suspended freely it will align itself so that one pole points roughly north and the other pole points roughly south. The pole that points north is called the north-seeking pole; the pole that points south is called the south-seeking pole. These names are now abbreviated to *north pole* and *south pole*.

A *natural magnet* is made by shaping a piece of magnetite so that the poles are at the ends. Early compasses were made using natural magnets. Today, natural magnets are not used, as better magnets can be made artificially. The method of making these *artificial magnets* will be described later in this chapter.

Magnetic poles

North and south poles always occur together in equal pairs. Such a pair of equal and opposite magnetic poles is called a *magnetic dipole*. An isolated north or south pole has never been observed. If a magnet is broken in two in an attempt to separate the north and south poles, new south and north poles appear, as shown in figure 8.6.

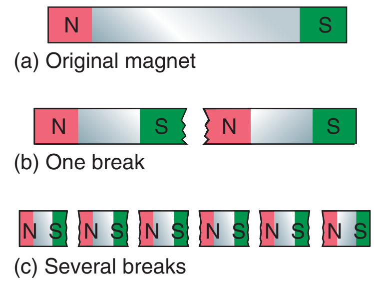


Figure 8.6 ***Breaking a bar magnet***

If two magnetic poles are brought close together they exert forces on one another as shown in figure 8.7. The directions of the forces between magnetic poles are:

- two north poles repel each other
- two south poles repel each other
- a north pole and a south pole attract each other.

That is, *like poles repel; unlike poles attract.*

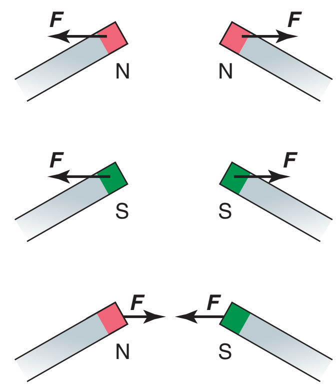


Figure 8.7 ***Forces between magnetic poles***

The closer two magnetic poles are to one another, the stronger the force of attraction or repulsion between them.

When two magnets are brought close to one another, there will be four pairs of forces between the poles. This will result in an overall force of attraction or repulsion (depending on the positions of the two magnets). An example of the forces between two magnets is shown in figure 8.8.

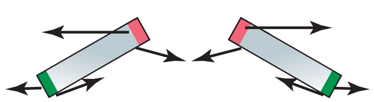


Figure 8.8 ***Forces between two magnets***

Magnetic fields

Interactions between magnetic poles can be described by magnetic fields, in the same way that interactions between electric charges were described by electric fields. In the field picture of magnetic interactions, each magnetic pole is surrounded by a **magnetic field** that exerts forces on other magnetic poles placed in the field. This is illustrated in figure 8.9.

The magnetic field is a force field surrounding a magnetic pole that exerts forces on other magnetic poles placed in the field.

The direction of a magnetic field is the direction of the force on a very small magnetic north pole placed in the field.

The magnetic field strength, B, which is studied in the HSC course, corresponds to the electric field strength, E.

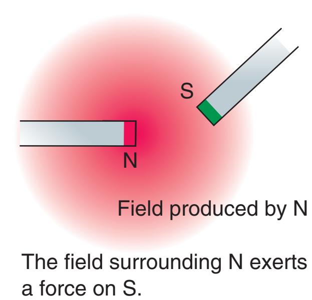


Figure 8.9 ***Magnetic field picture of magnetic interactions***

When a magnet is placed in a magnetic field, the north and south poles experience forces in opposite directions. The **direction of the magnetic field** at a point is defined as the direction of the force on a very small north pole placed at the point. The directions of the forces on the north and south poles of a magnet placed in a magnetic field are shown in figure 8.10. (In diagrams, a magnetic field is labelled ***B***.)

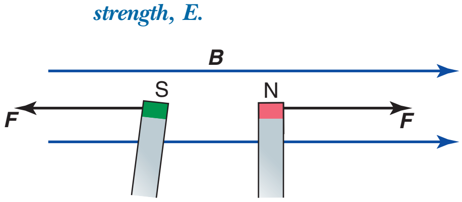
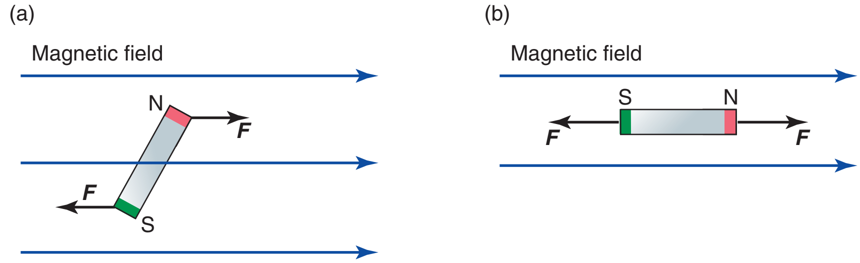


Figure 8.10 ***Forces on the north and south poles in a magnetic field***

A *compass* consists of a magnet suspended so that it is free to rotate. When a compass is placed in a magnetic field, the forces on the north and south poles cause the compass to rotate until the north pole of the compass points in the direction of the magnetic field. This is illustrated in figure 8.11.

Figure 8.11 *Forces on a compass in a magnetic field. (a) Forces causing compass needle to rotate and (b) compass needle aligned with magnetic field*



In practice, the direction of a magnetic field at a point is found by placing a small compass at the point. The direction in which the north pole of a compass points shows the direction of the magnetic field.

Magnetic field near a magnetic pole

Recall that the direction of the magnetic field at a point is the direction of the force on a very small north pole placed at the point.

If a compass is placed at a point near a north pole, N, the north pole of the compass will experience a force away from N and the south pole of the compass will experience a force towards N. The compass will point in the direction shown in figure 8.12a. The magnetic field therefore points away from N. Similarly, the magnetic field surrounding a south pole points towards the south pole. This is illustrated in figure 8.12b.

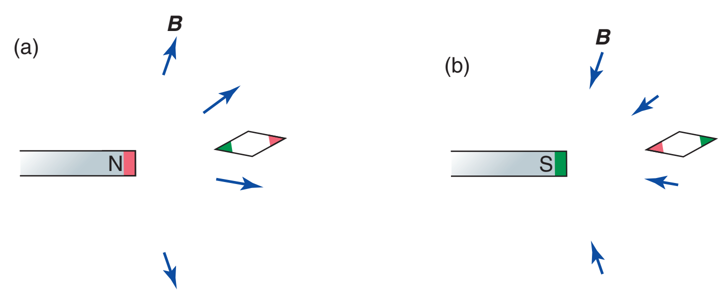


Figure 8.12 *(a) Magnetic field near a north pole (b) Magnetic field near a south pole*

Magnetic field surrounding two equal magnetic poles

- Magnetic fields are represented by magnetic field lines.
- Magnetic field lines start at north poles and end at south poles.
- The direction of the magnetic field lines shows the direction of the magnetic field.
- The spacing of the magnetic field lines shows the strength of the magnetic field. The closer the lines the stronger the field.

A magnet has equal north and south poles at the ends. The magnetic field surrounding a magnet is shown in figure 8.13.

Consider two identical magnets with their north poles placed close to one another. The magnetic field in the region near the two equal north poles is shown in figure 8.14a. Similarly, the magnetic field near the two equal south poles is shown in figure 8.14b.

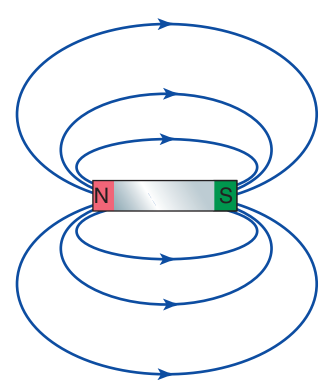


Figure 8.13 *Magnetic field surrounding a magnet*

8.2
Magnetic field surrounding a magnet

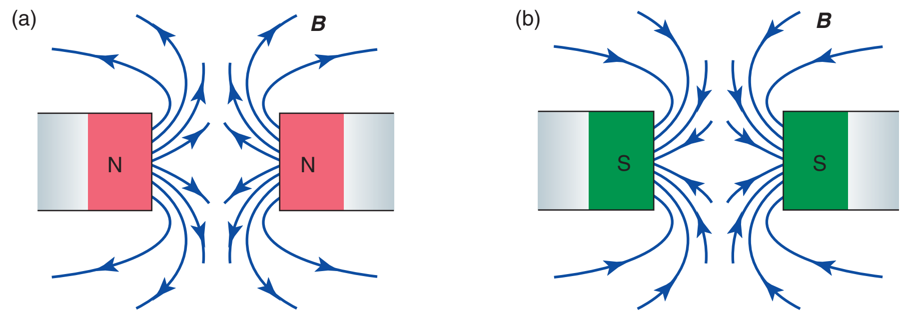


Figure 8.14 *(a) Magnetic field near two equal north poles (b) Magnetic field near two equal south poles*

8.3 MAGNETIC FIELDS AND ELECTRIC CURRENTS

Figure 8.16 *Hans Christian Oersted (1777–1851)*

Because the forces between magnetic poles are similar to the forces between electric charges, many early scientists suspected that there was a connection between magnetism and electricity.

In 1821, a Danish scientist, Hans Christian Oersted, while demonstrating to friends the flow of an electric current in a wire, noticed that the current caused a nearby compass needle to change direction (see figure 8.15). Oersted's observation showed that there was a magnetic field surrounding the electric current. Further investigation showed that all electric currents are surrounded by magnetic fields.

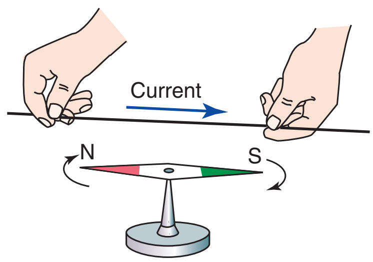


Figure 8.15 *Oersted's experiment*

Magnetic fields produced by electric currents

The magnetic field surrounding a wire carrying an electric current depends, in a complex way, on the shape of the wire. In this course, two important cases will be studied where the field produced by the current is comparatively simple.

Magnetic field surrounding a long, straight wire carrying a current

The magnetic field lines surrounding a long, straight wire carrying a current are concentric circles around the conductor. This is illustrated in figure 8.17.

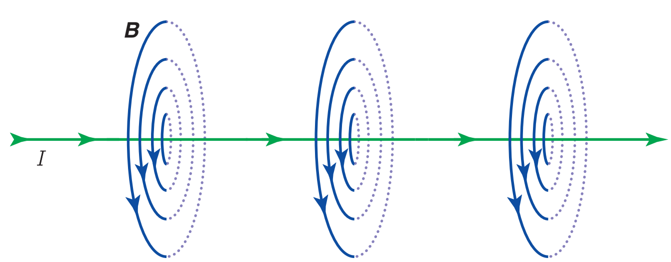


Figure 8.17 *Magnetic field surrounding a long, straight wire carrying a current*

The right-hand grip rule is a rule for finding the direction of the magnetic field surrounding an electric current.

The direction of the magnetic field is given by the **right-hand grip rule** (see figure 8.18). This states:

Grip the wire with the right hand, with the thumb pointing in the direction of the conventional current and the fingers will curl around the wire in the direction of the magnetic field.

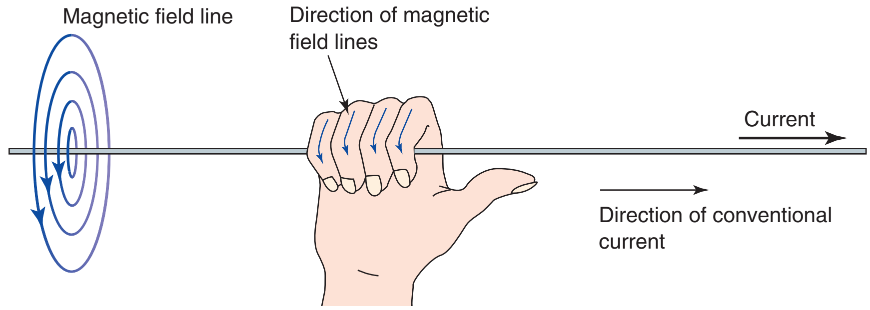
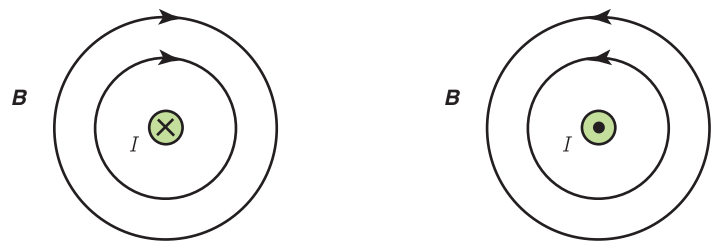


***Figure 8.18** Right-hand grip rule*

Recall that the direction of the conventional current is in the opposite direction to the direction of electron flow in the conductor.

In drawing a diagram to represent the magnetic field surrounding a wire carrying a current, it is often convenient to imagine the wire being perpendicular to the page. In such a diagram, the wire is represented by a small circle at the point where the wire passes through the page. The direction of the current will be *into the page* or *out of the page*. Figure 8.19 shows the magnetic fields surrounding electric currents passing into and out of the page.

The symbols '⊗' and '⊙' are used to represent directions into and out of the page. As an aid to remembering which is which, think of an arrow pointing into or out of the page. If the arrow is pointing out of the page, you will see the point of the arrow ⊙; if the arrow is pointing into the page, you will see the crossed feathers ⊗.



(a) Conventional current into page

(b) Conventional current out of page

***Figure 8.19** Magnetic fields surrounding currents passing into and out of the page*

PRACTICAL ACTIVITIES

8.3

Magnetic field produced by a current in a long, straight wire

Magnetic fields directed into and out of the page are represented by the symbols 'x' and '•'. Figure 8.20 shows the magnetic field at the sides of a wire carrying a conventional current towards the right of the page. By the right-hand grip rule, the magnetic field comes out of the page above the wire and goes into the page below the wire.

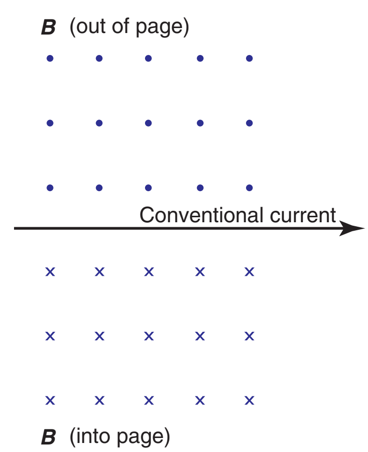


***Figure 8.20** Magnetic fields into and out of the page*

Magnetic field produced by a solenoid carrying a current

If a straight wire carrying a current is bent into a loop, the magnetic field is as shown in figure 8.21. The magnetic field lines come out at one side of the loop, which is therefore like the north pole of a magnet. The magnetic field lines go in to the other side of the loop, which is therefore like the south pole of a magnet. The right-hand grip rule, applied to a section of the loop, gives the direction of the magnetic field.

Magnetic fields produced by coils have many more practical applications than magnetic fields produced by straight wires.

The strength of the magnetic field can be increased by using a coil with many turns.

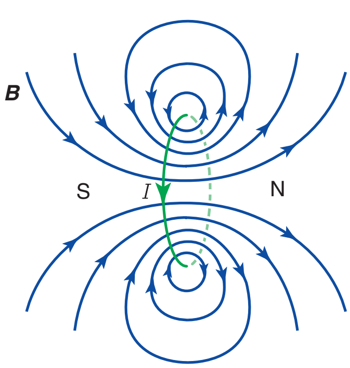


***Figure 8.21** Magnetic field produced by a current loop*

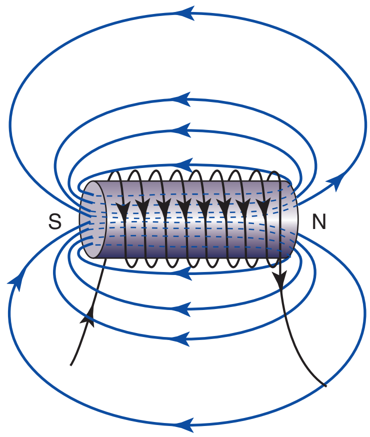


***Figure 8.22** Magnetic fields produced by a current in a solenoid*

A *solenoid* is a wire that has been wound into a closely packed helix (corkscrew shape). When a current passes through a solenoid, magnetic fields are produced both inside and outside the solenoid. The magnetic field outside the solenoid is similar to the magnetic field surrounding a bar magnet. For the solenoid, however, the lines of magnetic field do not stop at the ends of the solenoid but pass through the inside as parallel lines. The lines of magnetic field form closed loops. This is illustrated in figure 8.22. The end where the lines of magnetic field emerge from the solenoid is the north pole. The end where the lines of magnetic field enter the solenoid is the south pole. There are two methods of determining which end of a solenoid is the north pole.

1. Grip the solenoid with the right hand, with the fingers pointing in the direction of the conventional current around the solenoid. The thumb will point in the direction of the north pole of the solenoid. This is illustrated in figure 8.23.

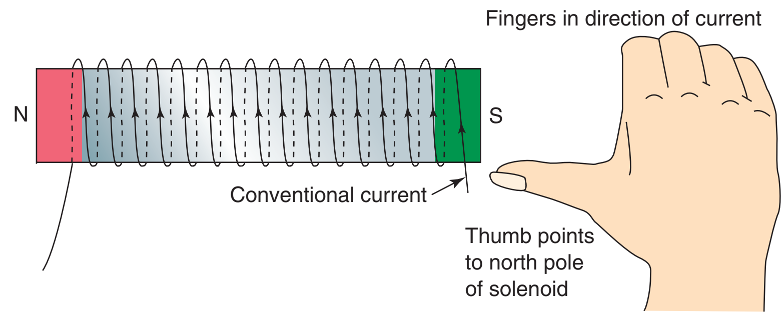
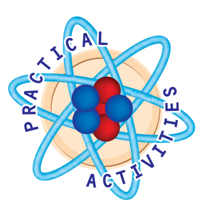


***Figure 8.23** Poles of a solenoid — first method*

2. Observe the solenoid end on. If the direction of the conventional current is anti-clockwise, the end is a north pole; if the conventional current is clockwise, the end is a south pole. This can be remembered by writing an N or an S with arrows as shown in figure 8.24.



8.4

Magnetic field of a solenoid carrying a current

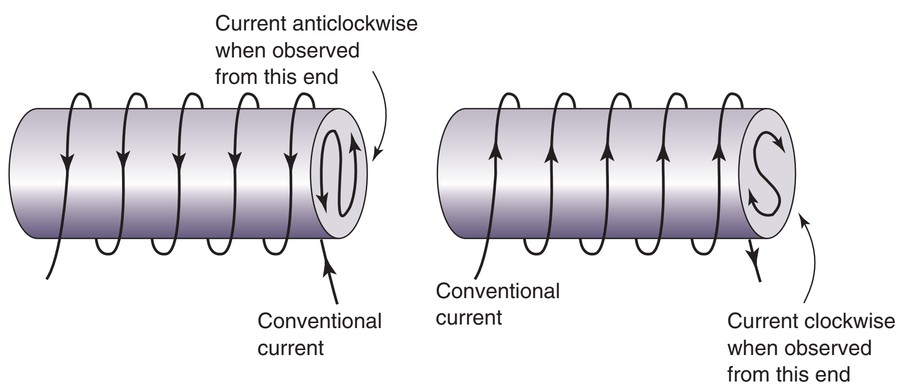


***Figure 8.24** Poles of a solenoid — second method*

Magnets and electromagnets

When a magnetic material, such as iron, is placed in a magnetic field, it becomes magnetised. 'Soft iron' is a type of iron that becomes magnetised very quickly when placed in a magnetic field, and loses its magnetism very quickly when removed from the field. When soft iron is magnetised by being placed in a magnetic field, it is said to be a *temporary magnet*.

A piece of iron is attracted to a magnetic pole because it becomes magnetised by the magnetic field surrounding the pole. Figure 8.25 shows a piece of iron being attracted by (a) a north pole and (b) a south pole.

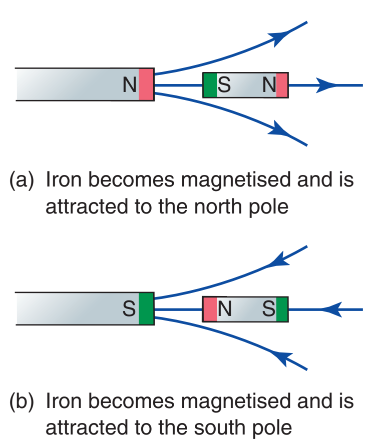
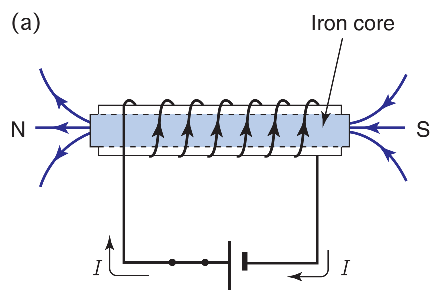


***Figure 8.25** Attraction of iron to (a) a north pole and (b) a south pole*

An alloy is a mixture of metals.

'Hard iron' is used to refer to any alloy of iron that becomes magnetised slowly when placed in a magnetic field, but retains its magnetism for a long time after it is removed from the field. Such an alloy is used to make a *permanent magnet*. To make a permanent magnet, a bar of hard iron is placed inside a solenoid, and a current is passed through the solenoid for a sufficient time to magnetise the iron.

A solenoid with a soft iron core is called an *electromagnet*. A current through the coil produces a magnetic field that magnetises the soft iron core almost instantaneously. This produces a much stronger magnet than would be produced by the solenoid without the soft iron core. When the current is switched off, the soft iron core loses its magnetism almost instantaneously. Figure 8.26 shows an electromagnet.



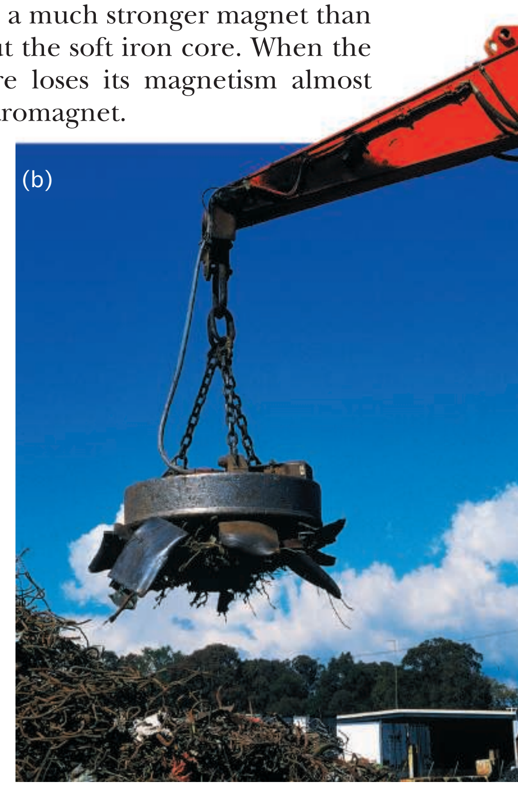


Figure 8.26 *(a) An electromagnet (b) An electromagnet in action*

Electromagnets have many practical applications. Figure 8.27 shows a circuit breaker that uses an electromagnet. If a sufficiently high current passes through the solenoid, the electromagnet will trip the mechanism and break the circuit.

PRACTICAL ACTIVITIES

8.5

Building an electromagnet

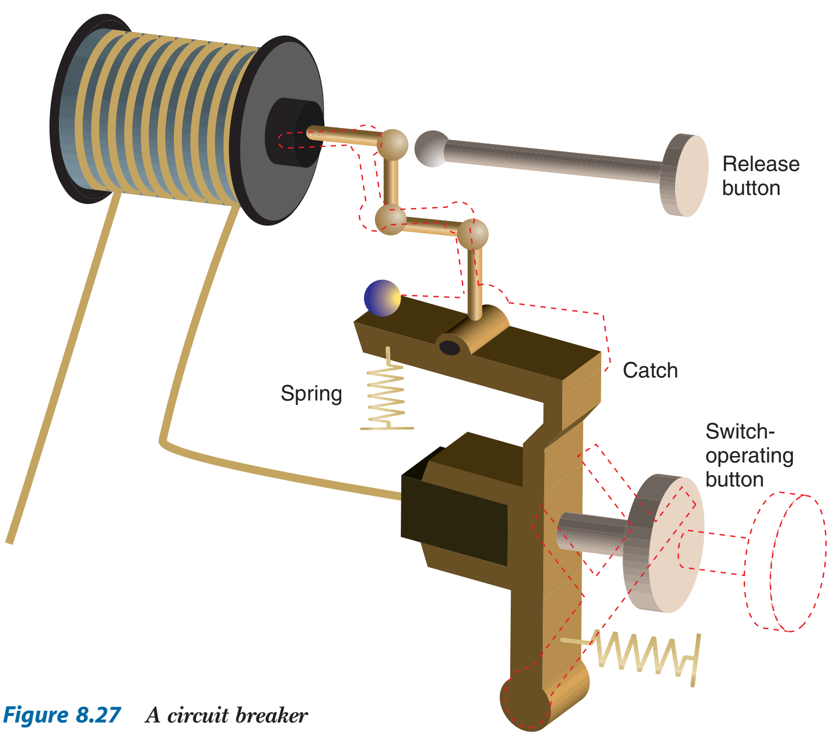


Figure 8.27 *A circuit breaker*

PHYSICS IN FOCUS

Ferromagnetism

It is now thought that *all* magnetic fields (including the magnetic fields of magnets) are produced by electric currents. This raises the question: 'Where is the electric current that produces the magnetic field of a magnet?'.

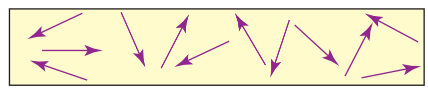
Within every atom there are electric currents due to the movement of the electrons. Each electron moves in an orbit around the nucleus and spins on its axis. (This picture of the motion of an electron in an atom is much simpler than what physicists believe to be the actual situation.)

As a result of the movement of its electrons, each atom behaves like a current loop with a north and south pole. The magnetic properties of a material are due to the magnetic properties of its atoms.

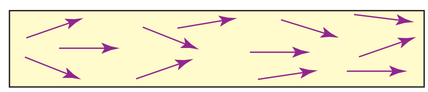
As all materials consist of atoms, it follows that all materials will have magnetic properties. For most materials, the magnetism is very weak and cannot be detected without using very strong magnetic fields. In a few materials, the magnetism is strong. In these materials, the spins of the electrons producing the atomic magnetism tend to line up in neighbouring atoms. As a result, the north and south poles of the atomic magnets tend to line up and point in the same direction producing a strong magnetic field. These materials are called ferromagnetic materials. Examples of ferromagnetic materials are iron, cobalt and nickel. Magnets are made of ferromagnetic materials.

A ferromagnetic material is made up of many *magnetic domains*. Within each domain nearly all the atomic magnets are lined up in the same direction. Each domain contains 10^9–10^{15} atoms. When a ferromagnetic material is unmagnetised, the domains are magnetised in random directions so that there is no overall magnetism.

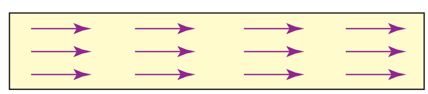
When the ferromagnetic material is placed in a magnetic field, the directions of magnetism of the domains tend to line up in the direction of the magnetic field. When all the domains are lined up with the magnetic field, the material is fully magnetised, or saturated. Figure 8.28 shows the magnetic domains in an unmagnetised, a partially magnetised and a fully magnetised sample of ferromagnetic material.



(a) Random direction of domains of an unmagnetised magnetic substance



(b) Domains begin to face in same direction, producing a weak magnet.



(c) All domains face the same direction, resulting in a strong magnet.

Figure 8.28 ***Magnetic domains***

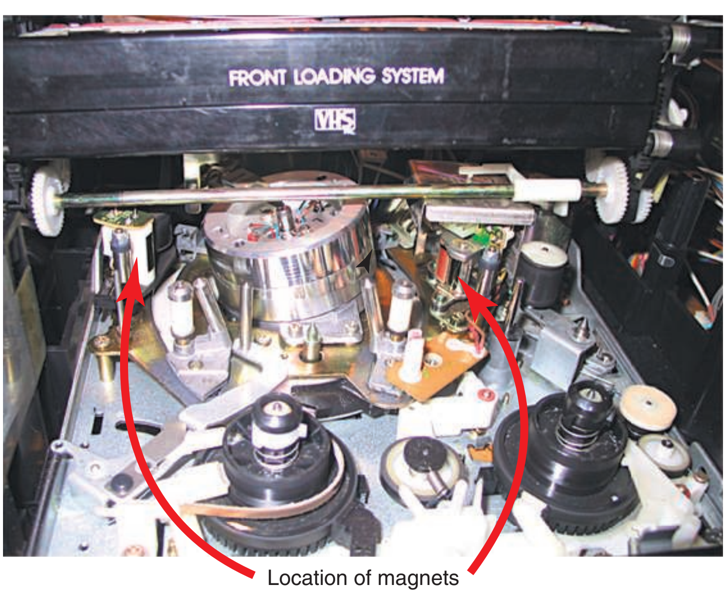


Figure 8.29 ***A video tape recorder is an example of a household appliance that uses magnetic fields.***

Tape recorders — an application of magnetic fields

Magnetic tape consists of a coating of microscopic particles of magnetic material suspended on a plastic backing such as acetate.

Figure 8.30 shows the recording head of a recorder. During recording, an alternating current, derived from sound waves, is passed through the coil of the recording head. This produces a varying magnetic field in the air gap between the poles that magnetises the magnetic particles on the tape.

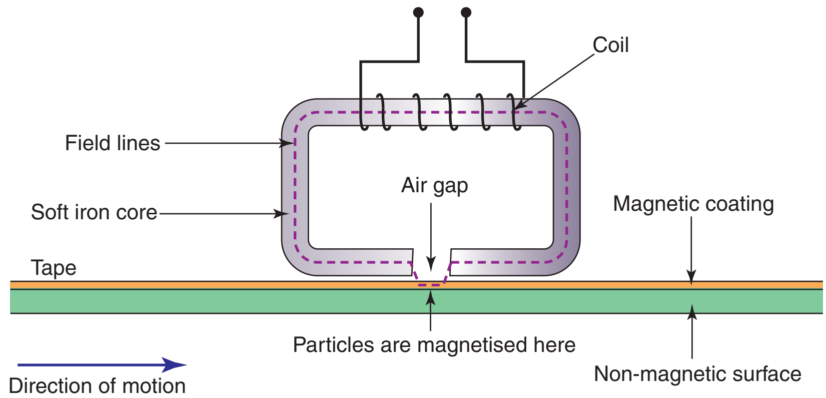


Figure 8.30 *Recording on a magnetic tape*

Induced currents are studied in the HSC course.

Figure 8.31 shows how the direction of magnetisation of the particles corresponds with the direction of the input signal. Later, during playback, these magnetised particles induce an alternating current which reproduces the original alternating current.

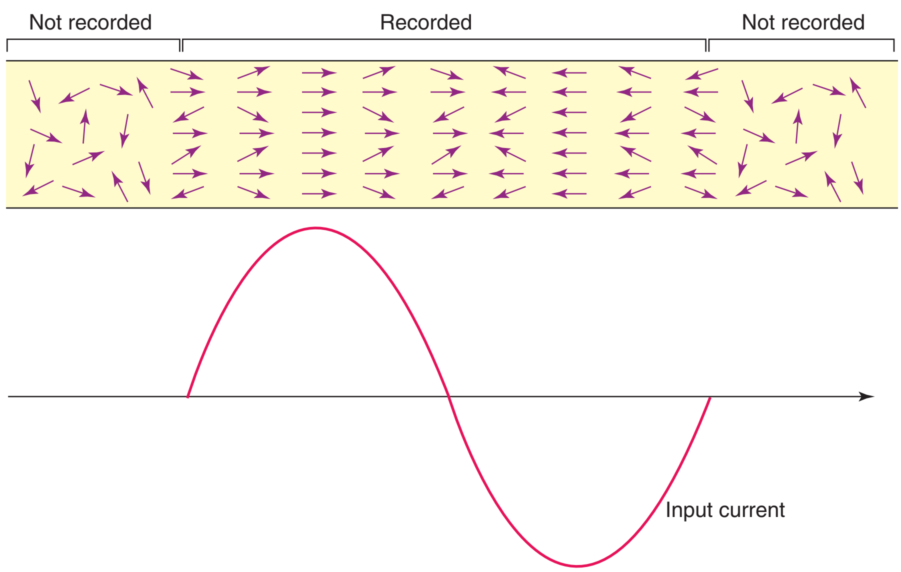


Figure 8.31 *Magnetisation of recording tape*

SUMMARY

- Power is the rate at which energy is transformed from one form to another.
- $P = \dfrac{W}{t}$
- $P = VI$
- A kilowatt-hour (kW-h) is a unit of energy equal to the amount of energy used by a 1 kW device in 1 hour. As it is much larger than 1 joule, it is a more convenient unit in which to measure household energy consumption.
- $W = VIt$
- Like magnetic poles repel; unlike magnetic poles attract.
- The direction of a magnetic field is the direction of the force that acts on a very small north pole placed in the field.
- Magnetic fields point away from north poles and towards south poles.
- Magnetic field lines are used to represent magnetic fields.
- An electric current in a long, straight wire produces a magnetic field represented by field lines in the form of concentric circles around the wire.
- The right-hand grip rule relates the direction of the current in a wire to the direction of the magnetic field.
- The magnetic field produced by a current in a solenoid is similar to that produced by a magnet.
- An electromagnet consists of a solenoid with a soft iron core. When a current flows through the solenoid the soft iron core becomes strongly magnetised.
- Tape recorders are household appliances that use magnetic fields.

QUESTIONS

1. Calculate the power generated by a battery that produces 6.00×10^2 J of electric potential energy in 2.00×10^1 s.
2. Calculate the heat energy produced in 1.00 minute by a heater generating 5.00×10^2 W of power.
3. Calculate the time it takes a 4.00×10^1 W lamp to dissipate 8.00×10^2 J of electrical energy.
4. Calculate the electrical energy (in kW-h) used by a 3.0 kW heater in 2.0 hours.
5. The following table refers to the energy, *W*, produced by a device which generates power, *P*, for time *t*. Calculate the missing values and complete the table.

	P	*W*	*t*
(a)	5.0 kW		2.0 min
(b)		2.00×10^4 J	1.60×10^3 s
(c)	6.2×10^4 W		1.0 h
(d)	7.5 MW	2.5 kJ	
(e)		2.40×10^2 J	35.0 s
(f)	2.0×10^3 W	2.5×10^4 J	
(g)	2.8 kW		1.5 s
(h)		2.8×10^1 J	1.4×10^{-1} s
(i)	1.20 kW		4.50 h
(j)		14.4 kW-h	7.20 h
(k)	2.50 kW	8.25 kW-h	

6. The following table refers to the power, *P*, generated when a current, *I*, passes through a potential difference (voltage), *V*. Calculate the missing values and complete the table.

	P (W)	*I* (A)	*V* (V)
(a)	4.20	2.00	
(b)	10.0		5.00
(c)		2.50	4.00
(d)	2.0		6.0
(e)	6.8	13.6	
(f)		12.5	5.00

7. *W* is the energy dissipated when a current, *I*, passes for a time, *t*, through a resistor with a voltage drop, *V*, across it. Calculate the missing values and complete the table below.

	W (J)	*I* (A)	*V* (V)	*t* (s)
(a)	24.0		3.00	4.00
(b)	50.0	10.00	2.00	
(c)	60.0	12.0		5.00
(d)		2.00	5.26	20.4
(e)	1.71×10^3		12.0	25.3
(f)		2.30×10^{-2}	25.0	2.77×10^2
(g)	1.08×10^6	2.50	2.40×10^2	
(h)	6.86×10^4	5.25		3.75×10^2

8. The following table refers to the power, P, generated when a current, I, passes through a resistor of resistance R with a potential difference (voltage), V, across it. Calculate the missing values and complete the table.

	POWER (W)	CURRENT (A)	VOLTAGE (V)	RESISTANCE (Ω)
(a)	1.0×10^3		2.40×10^2	
(b)		3.20	1.12×10^2	
(c)		2.0		48
(d)			16	2.0
(e)	1.0×10^2	1.5×10^{-1}		
(f)	5.00×10^2		2.00×10^2	

9. Two identical magnets are placed in each of the positions shown in figure 8.32. For each case draw a diagram showing the forces acting between the poles of one magnet and the poles of the other magnet (four pairs of forces in each case). Indicate the strengths of the forces by the lengths of the lines representing the forces. For each case, state whether the magnets will attract or repel one another.

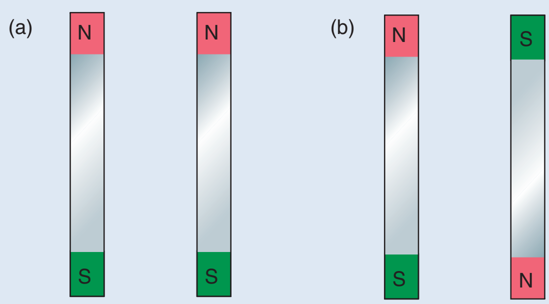


Figure 8.32

10. In each of the cases shown in figure 8.33, sketch the magnetic field and mark the direction of the magnetic field at the points X, Y and Z.

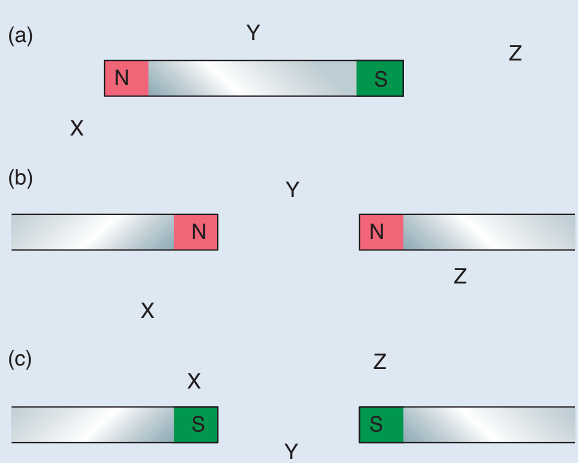


Figure 8.33

11. Sketch the magnetic fields due to the solenoids shown in figure 8.34. Mark the north and south poles in each case. Show the direction in which a compass needle will point at points X, Y and Z in each case. In each case, state the direction in which a compass needle inside the solenoid would point.

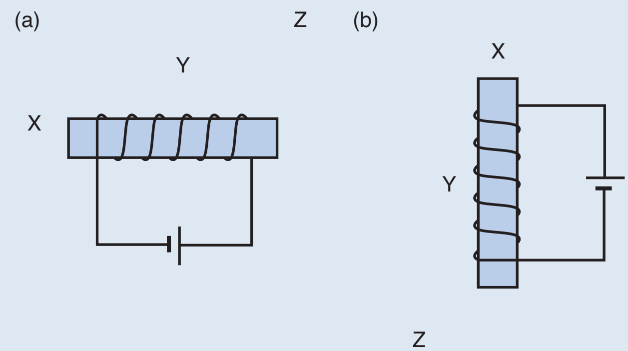


Figure 8.34

12. The diagrams in figure 8.35 represent currents in wires perpendicular to the page (⊗ represents a current into the page and ⊙ represents a current out of the page). For each case, draw lines of magnetic field strength to represent the magnetic field surrounding the wire. Mark the directions in which the north pole of a compass would point at X, Y and Z in each of the cases shown.

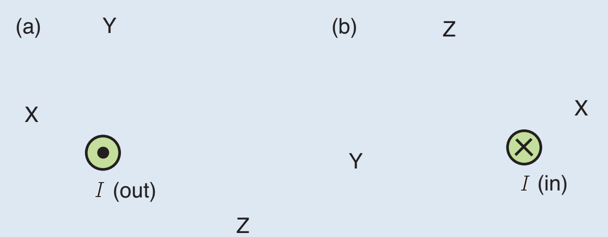


Figure 8.35

13. In each of the cases shown in figure 8.36, show the magnetic field using the symbol 'x' to represent a field into the page and the symbol '•' to represent a field out of the page.

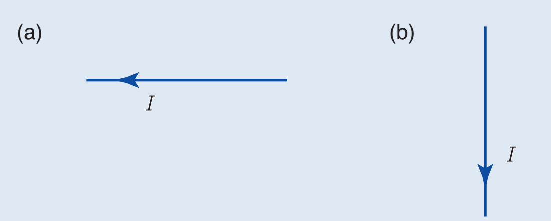


Figure 8.36

14. Figure 8.37 represents a magnet placed in a magnetic field. Draw a diagram showing the forces that act on the poles of the magnet. If the magnet is free to move, how will it move? Justify your answer.

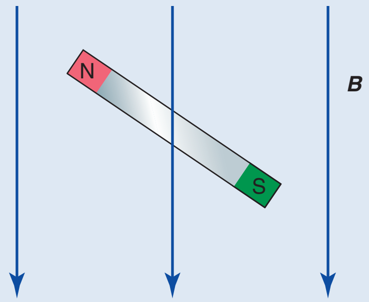


Figure 8.37

15. Figure 8.38 shows two electromagnets. Will they attract or repel one another? Justify your answer.

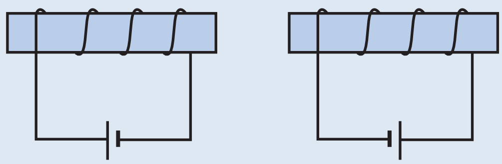


Figure 8.38

16. Explain how magnetic fields are used in a tape recorder.
17. Refer to the electricity account on page 153.
    (a) In what units is the energy measured?
    (b) What is the normal cost per unit of energy?
    (c) What is the off-peak cost per unit of energy?
    (d) What would the hot water cost if it was charged at the normal rate?
    (e) What was the average amount of energy used by this household per day?

8.1 MODEL HEATING COIL

Aim

To design, make and test a 6 V to 12 V heating coil which will raise 200 mL of water from room temperature to boiling point in approximately 10 minutes

Apparatus

power supply
variable resistor
graduated beaker
ammeter
voltmeter
switch
connecting wires
thermometer
resistance wire of known resistance per centimetre
centimetre ruler

Theory

1. The energy (W) needed to raise the temperature of water is given by:
   W = (mass of water in grams) × 4.2 × (change in temperature in °C).
2. The power required is given by the formula:
   $P = \frac{W}{t}$.
3. The resistance needed to generate this power can be calculated using the formula $R = \frac{P}{I^2}$.
4. 1 mL of water has a mass of 1 g.

Method

1. Find out from your teacher what current you should use.
2. Place 200 mL of water from the tap into a beaker and measure its temperature.
3. Using the theory above, calculate:
   (a) the amount of energy required to raise the temperature of the water to 100°C
   (b) the power which must be dissipated to raise the temperature of the water to 100°C in 10 minutes
   (c) the resistance of the coil needed to provide this power
   (d) the length of resistance wire required.
4. Obtain this length of resistance wire and make a coil suitable for heating the water in your beaker.
5. Suspend the coil so that it is immersed in the water. Do not let it touch the sides of the beaker.
6. Connect your apparatus as shown in figure 8.39.

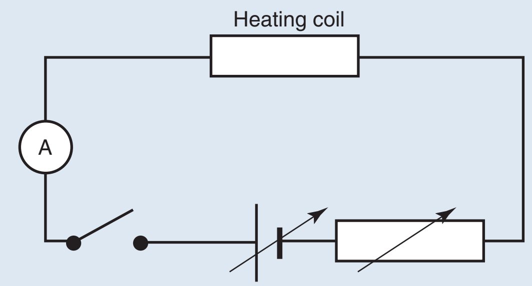


Figure 8.39 *Circuit diagram*

7. Use the power supply and the variable resistor to set the current to the required value.
8. Test your coil by heating 200 mL of water and measuring the time taken to raise the temperature to boiling point.

Results

Record your results as:
Temperature of water =
Rise in temperature of the water =
Energy required to raise temperature of water to 100°C =
Power to raise temperature in 10 minutes =
Current =
Resistance of coil =
Resistance per centimetre of resistance wire =
Length of resistance wire required =
Time taken to raise temperature of water =

Analysis

1. Did it take more or less time to raise the temperature of the water than you had predicted?
2. What was the difference in time between the predicted value and the measured value?
3. Express this as a percentage of the actual value.
4. What assumptions were made in your calculations that may have been incorrect?

Question

How would you improve the accuracy of this activity?

8.2 MAGNETIC FIELD SURROUNDING A MAGNET

Aim

To use a compass to map the magnetic field surrounding a bar magnet

Apparatus

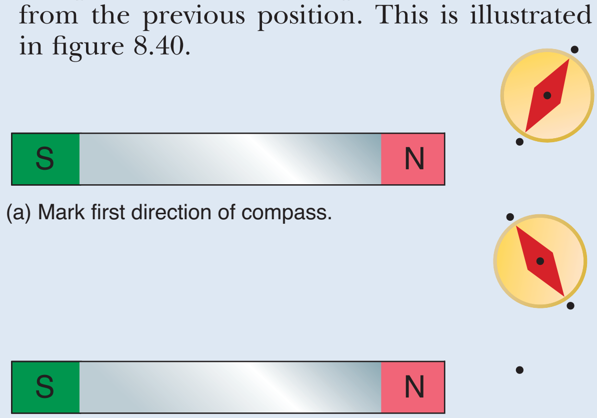
bar magnet
compass
large sheet of paper

Theory

The direction of a magnetic field at a point can be found by placing a small compass at the point. The north pole of the compass points in the direction of the magnetic field at the point.

Method

1. Place the sheet of paper on a horizontal surface.
2. Use the compass to find the N–S direction and mark this direction at the centre of the paper.
3. Place the bar magnet on the paper along the N–S line marked on the paper with the north pole of the magnet pointing north.
4. Mark on the paper the outline of the magnet and label the poles N and S.
5. Place the compass at a point near the north pole of the magnet. Mark with two points the position taken up by the compass needle.
6. Move the compass to a new position so that the position of the compass needle follows from the previous position. This is illustrated in figure 8.40.



(a) Mark first direction of compass.

(b) Place compass so that the direction follows on from the previous direction.

Figure 8.40

7. Continue in this way until you reach a position near the south pole of the magnet.
8. Draw a continuous curve through the points you have marked on the paper.
9. Mark with arrows the direction of the magnetic field at several points along your line.
10. Repeat five times starting from different positions of the compass.

Analysis

1. At which pole do the magnetic field lines begin?
2. At which pole do the magnetic field lines end?
3. Where is the magnetic field strongest? How is this shown by the magnetic field lines?

8.3 MAGNETIC FIELD PRODUCED BY A CURRENT IN A LONG, STRAIGHT WIRE

Aim

To map the magnetic field surrounding a long, straight wire carrying an electric current

Apparatus

50 cm length of straight wire
sheet of cardboard approximately 20 cm × 20 cm
power supply
variable resistor
connecting wire
switch
compass
some means of supporting the wire
some means of supporting the cardboard

Method

1. Set up the apparatus as shown in figure 8.41. To increase the strength of the magnetic field, a number of loops of wire can be used.

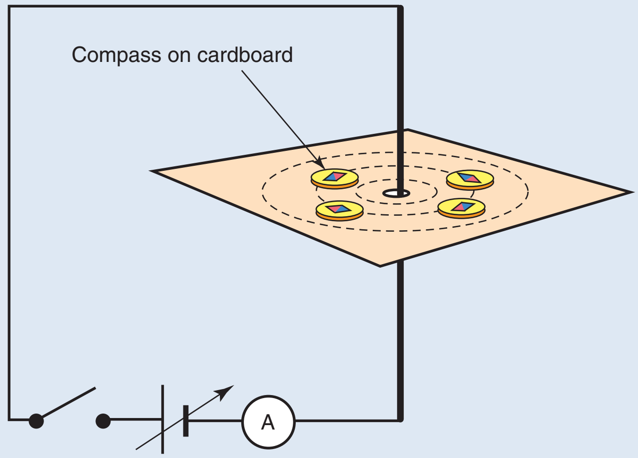


Figure 8.41

2. Connect the power supply so that the conventional current flows downwards through the wire.
3. Adjust the voltage of the power supply and the variable resistance so that the current has the value given by your teacher.
4. Place the compass about 5 cm from the wire.
5. Switch on the current and mark the positions of the ends of the compass on the cardboard.
6. Proceed as in practical activity 8.2, tracing out the magnetic field line. (Ideally this should return to the starting point to form a closed loop.)
7. Mark the direction in which the north pole of the compass pointed at several places on the magnetic field line.
8. Repeat this a number of times with the initial position of the compass at different distances from the wire.
9. Draw smooth lines of magnetic field through each set of points.
10. Reverse the direction of the current and observe what happens to the compass needle.

Analysis

Show that your result is compatible with the right-hand grip rule.

Questions

1. When the current was coming upwards out of the cardboard, was the direction of the magnetic field lines around the wire clockwise or anticlockwise?
2. Can you use this to formulate an alternative rule for determining the direction of the magnetic field surrounding a current-carrying wire?

8.4 MAGNETIC FIELD OF A SOLENOID CARRYING A CURRENT

Aim

To map the magnetic field surrounding a solenoid

Apparatus

solenoid
sheet of cardboard approximately 20 cm × 20 cm
scissors
connecting wires
power supply
variable resistor
switch
compass

Method

1. Connect the apparatus as shown in figure 8.42.

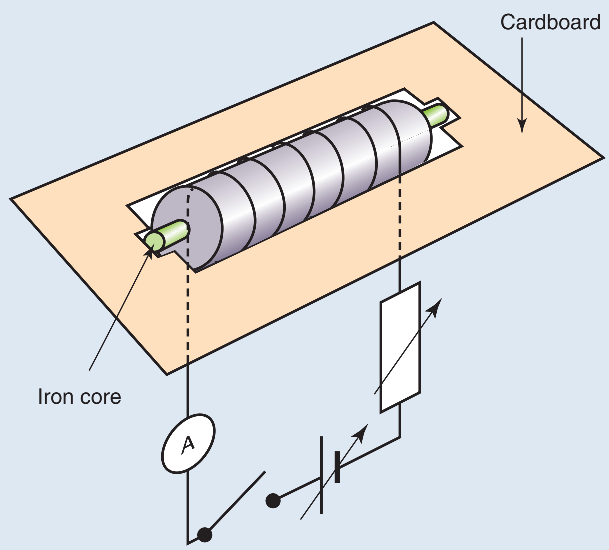


Figure 8.42

2. Note the direction of the conventional current around the solenoid.
3. Map the magnetic field around the solenoid using the same method as was used in the previous two practical activities.
4. Reverse the direction of the current through the solenoid. Note what happens to the direction of the magnetic field.

Analysis

1. With the first direction of the conventional current, which end of the solenoid was the north pole? Explain.
2. Is this result compatible with the right-hand grip rule for solenoids?
3. Draw a sketch showing how the right-hand grip rule for solenoids applies to your result.
4. What happened to the magnetic field when the direction of the current was reversed?
5. How does the magnetic field produced by a current in a solenoid compare with the magnetic field surrounding a magnet?

8.5 BUILDING AN ELECTROMAGNET

Aim

To build an electromagnet and observe its properties

Apparatus

iron rod for core of electromagnet
insulated conducting wire for coil
power pack
connecting wire
variable resistor
ammeter
small iron nails

Theory

A soft iron core is placed in a solenoid carrying a current and becomes magnetised. When the current is switched off the soft iron core loses its magnetism.

Method

1. Build the electromagnet by winding the conducting wire closely from one end of the iron core to the other. To make the electromagnet stronger one or more layers of coils can be wound on top of the first. It is essential that all layers of coils are wound in the same direction around the core.
2. Connect the electromagnet to the power supply as shown in figure 8.43.

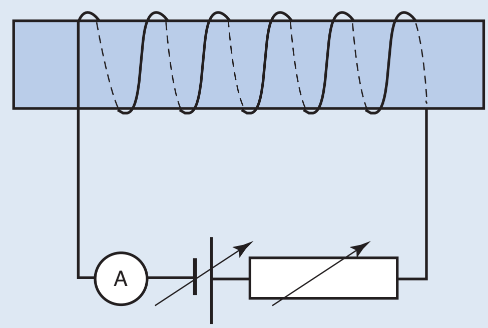


Figure 8.43

3. Test the magnetism of the electromagnet by observing the attraction of small iron nails to the end of the soft iron rod. The greater the number of iron nails attracted to the rod, the greater is the magnetism.
4. Observe the magnetism of the electromagnet when there is no current.
5. Observe the magnetism of the electromagnet for a range of currents.
6. Observe how much time is taken for the electromagnet to gain and lose its magnetism when the current is switched on and off.

Analysis

1. How did the magnetism change as the current was increased?
2. Was there any delay observed in the gain or loss of magnetism when the current was switched on or off?
3. Was there any magnetism left when the current was turned off?

Chapter 9
Describing movement

Chapter 10
Force and Newton's laws of motion

Chapter 11
Mechanical interactions

PRELIMINARY MODULE

MOVING ABOUT

CHAPTER 9 DESCRIBING MOVEMENT

Figure 9.1 *Whether you're riding a bike, driving a car or bus, or even flying a plane, you need to be able to describe your movement in terms of your position, speed, direction and acceleration.*

Remember

Before beginning this chapter, you should be able to:

- describe the relationship between distance, speed and time
- relate acceleration to change in speed and/or direction of an object over a period of time.

Key content

At the end of this chapter you should be able to:

- distinguish between scalar and vector quantities
- identify changes in speed and velocity during a journey
- compare the instantaneous speed and instantaneous velocity of moving objects
- distinguish between instantaneous and average speed, and between instantaneous and average velocity
- describe the speed of one vehicle relative to another vehicle
- define average velocity as $\frac{\Delta \boldsymbol{r}}{\Delta t}$
- define average acceleration as $\frac{\Delta \boldsymbol{v}}{\Delta t}$ and therefore $\frac{\boldsymbol{v} - \boldsymbol{u}}{t}$
- use vector diagrams to determine displacement, change in velocity, and acceleration
- use equations relating a, v, u, r and t to make predictions about the motion of objects
- use graphical methods to record and make predictions about the motion of objects.

Vector quantities can be described in writing or by labelled arrows. If a symbol is used to represent a vector quantity, it may have a half-arrow above it or a 'squiggly' line below it. In this text, vector quantities are represented by symbols in bold italic type.

Most people today rely on some form of transport to get to school or work and to get around on weekends or during holidays. Whether you ride, drive, fly or sail, you need to know how far you are going, in which direction and when you intend to arrive. Whether or not you arrive on time depends on how fast you move and the direction you take. Describing motion is important in planning a journey, even if it is by foot. The study of motion is called kinematics.

9.1 DISTANCE AND DISPLACEMENT

A scalar quantity specifies size (magnitude) but not direction.

Displacement is a measure of the change in position of an object. It is a vector quantity.

A vector quantity specifies size (magnitude) and direction.

Distance is a measure of the total length of the path taken during the change in position of an object. Distance is a **scalar** quantity. Scalar quantities are those that specify size or magnitude, but not direction. Other examples of scalar quantities include speed, time and energy.

Displacement is a measure of the change in position of an object. Displacement is a **vector** quantity. Vector quantities are those that specify a direction as well as a size. In order to fully describe a displacement, a direction must be specified as well as a magnitude. The symbol $\boldsymbol{r}$ is used to represent position. Displacement is therefore represented by $\Delta\boldsymbol{r}$.

SAMPLE PROBLEM 9.1

Distance and displacement

A cyclist rides a distance of 6 km in a northerly direction, turns right and travels 8 km in an easterly direction. Determine:

(a) the distance travelled by the cyclist

(b) the displacement of the cyclist.

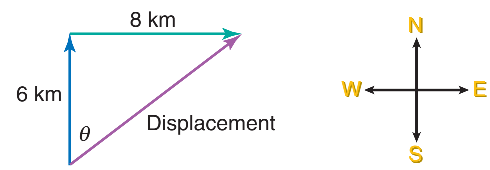


Figure 9.2 *The path taken by the cyclist*

SOLUTION

(a) The distance travelled = 6 + 8
= 14 km

(b) The magnitude of the displacement is given by:

$$r = \sqrt{(6)^2 + (8)^2}$$
$$= \sqrt{100}$$
$$= 10 \text{ km.}$$

The direction of the displacement is given by:

$$\tan\theta = \frac{8}{6}$$
$$\theta = 53.1°.$$

The displacement of the cyclist is 10 km north 53.1° east.

Adding vectors

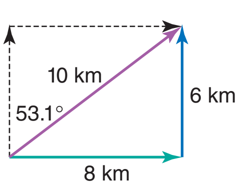


Figure 9.3 *The order in which vectors are added doesn't affect the result.*

The total displacement of the cyclist in sample problem 9.1 is a measure of the cyclist's change in position. However, it is also the sum of two other displacement vectors. When vector quantities like displacement are added, the labelled arrows that represent the vectors are placed 'head to tail'. The sum of the vectors is represented by the arrow drawn from the tail of the first vector and the head of the second vector. The order in which the vectors are added does not matter (see figure 9.3). For example, if the cyclist travelled 8 km east before travelling 6 km north the sum of the displacement vectors would be the same; that is, 10 km north 53.1° east.

SAMPLE PROBLEM 9.2

Vector addition

Add the following pairs of vectors to find the total displacement:
(a) displacement of 40 m east, displacement of 30 m east
(b) displacement of 40 m east, displacement of 30 m west
(c) displacement of 40 m east, displacement of 30 m north.

SOLUTION

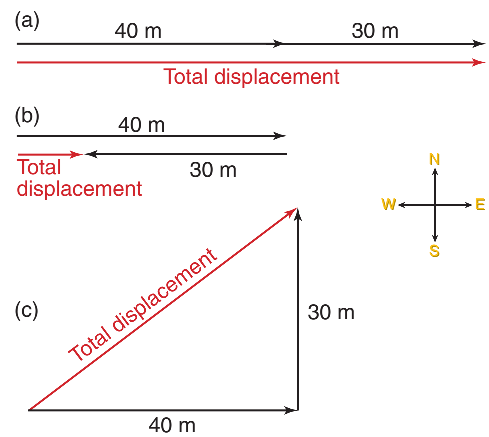


Figure 9.4

9.2 SPEED AND VELOCITY

Speed is a measure of the time rate at which an object moves over a distance. It is a scalar quantity.

Speed is a measure of the rate at which an object moves over a distance.

The average speed of an object can be calculated by dividing the distance travelled by the time taken, that is:

$$\text{average speed} = \frac{\text{distance travelled}}{\text{time taken}}.$$

Speed is a scalar quantity. The unit of speed is m s^{-1} if SI units are used for distance and time. However, it is often more convenient to use other units such as cm s^{-1} or km h^{-1}.

PHYSICS FACT

A snail would lose a race with a giant tortoise! A giant tortoise can reach a top speed of 0.37 km h^{-1}. However, its 'cruising' speed is about 0.27 km h^{-1}. The world's fastest snails cover ground at the breathtaking speed of about 0.05 km h^{-1}. However, the common garden snail is more likely to move at a speed of about 0.02 km h^{-1}. Both of these creatures are slow compared with light, which travels through the air at 1080 million km h^{-1}, and sound, which travels through the air (at sea level) at about 1200 km h^{-1}.

How long would it take the snail, giant tortoise, light and sound respectively to travel once around the equator, a distance of 40 074 km?

Velocity is a measure of the time rate of displacement, or the time rate of change in position. It is a vector quantity.

In everyday language, the word **velocity** is often used to mean the same thing as speed. In fact, velocity is not the same quantity as speed. Velocity is a measure of the time rate of displacement, or the time rate of change in position. Because displacement is a vector quantity, velocity is also a vector quantity. The velocity has the same direction as the displacement. The symbol $\boldsymbol{v}$ is used to denote velocity.

The average velocity of an object, $\boldsymbol{v}_{av}$, during a time interval, Δt, can be expressed as:

$$\boldsymbol{v}_{av} = \frac{\Delta \boldsymbol{r}}{\Delta t}$$

where, $\Delta \boldsymbol{r}$ = displacement (change in position).

For motion in a straight line in one direction only, the magnitude of the velocity is the same as the speed. The motion of the cyclist in sample problem 9.1 illustrates the difference between velocity and speed. If the trip takes 20 minutes, the cyclist's average velocity is:

$$\boldsymbol{v}_{av} = \frac{\Delta \boldsymbol{r}}{\Delta t} = \frac{10 \text{ km north } 53.1° \text{ east}}{20 \text{ minutes}}$$

$$= 0.50 \text{ km min}^{-1} \text{ north } 53.1° \text{ east.}$$

PRACTICAL ACTIVITIES

9.1

Going home

However, the average speed of the cyclist is:

$$\text{average speed} = \frac{\text{distance travelled}}{\text{time taken}}$$

$$= \frac{14 \text{ km}}{20 \text{ min}}$$

$$= 0.70 \text{ km min}^{-1}.$$

SAMPLE PROBLEM 9.3

Speed and velocity of a swimmer

If champion swimmer Grant Hackett completes 30 laps of a 50 m long swimming pool, a distance of 1500 m, in a time of 15 minutes, what is:

(a) his average speed in m s^{-1}
(b) his average velocity in m s^{-1}?

SOLUTION

(a) $\text{Average speed} = \frac{\text{distance travelled}}{\text{time taken}}$

$$= \frac{1500}{15 \times 60}$$

$$= 1.7 \text{ m s}^{-1}$$

(b) $\text{Average velocity} = \frac{\Delta \boldsymbol{r}}{\Delta t}$

$$= \frac{0}{15 \times 60}$$

$$= 0 \text{ m s}^{-1}$$

It's all relative

The velocity of an object measured by a moving observer is referred to as the relative velocity. The relative velocity is the difference between the velocity of the object relative to the ground and the velocity of the observer relative to the ground.

Imagine that you are in a car travelling at a constant velocity of 90 km h^{-1} due west on a straight road. The car ahead of you is travelling at a constant speed of 100 km h^{-1} in the same direction. Although the velocity of the other car relative to the road is 100 km h^{-1} due west, its velocity relative to you is 10 km h^{-1} due west. That is, the velocity of the car relative to you is equal to 100 km h^{-1} due west (velocity of car relative to the ground) minus 90 km h^{-1} due west (your velocity relative to the ground): 10 km h^{-1} due west. This is illustrated in figure 9.5a.

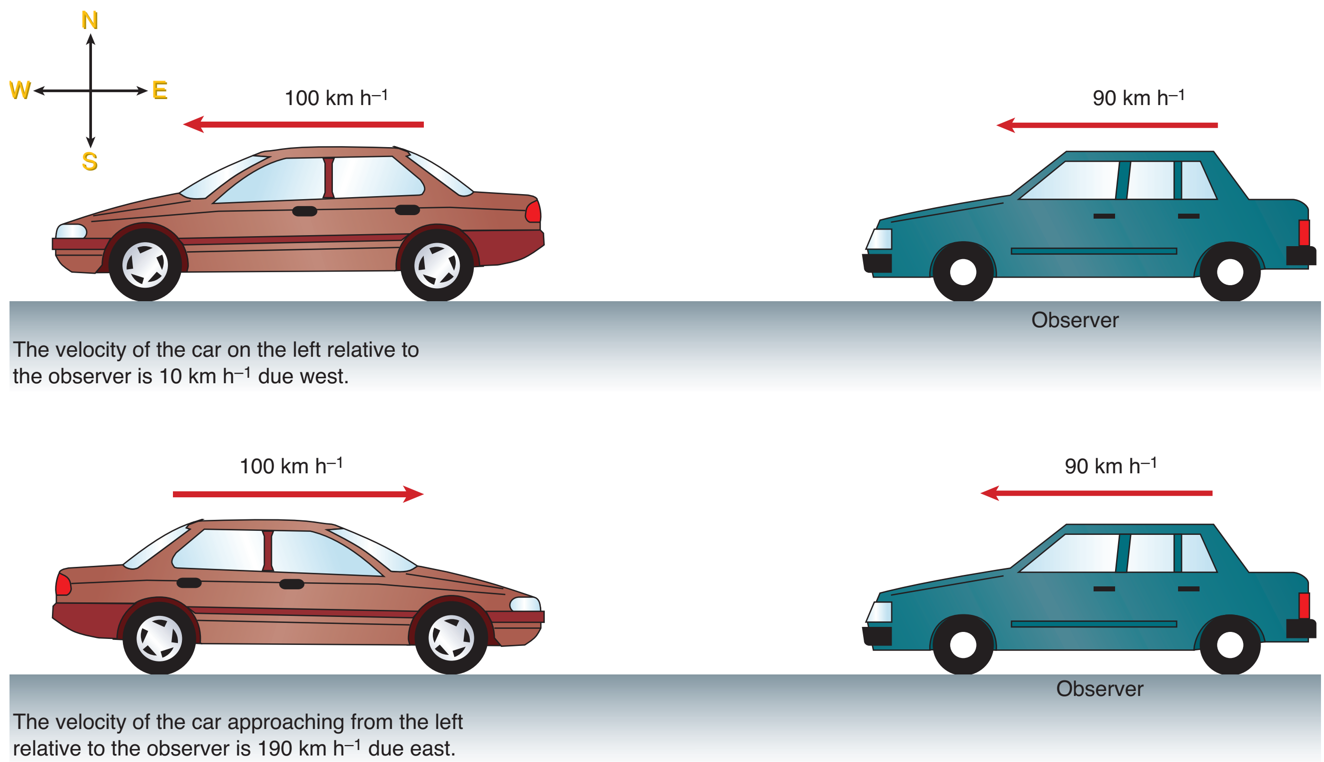


***Figure 9.5** The velocity that is measured depends on the velocity of the observer.*

If another vehicle were approaching you at a speed of 100 km h^{-1} relative to the road; that is, with a velocity of 100 km h^{-1} due east relative to the road, its velocity relative to you would be the difference between 100 km h^{-1} due east and 90 km h^{-1} due west. A velocity of 90 km h^{-1} due west is the same as –90 km h^{-1} due east. The relative velocity is therefore 100 km h^{-1} due east (velocity of car relative to the ground) minus –90 km h^{-1} due east (your velocity relative to the ground): 190 km h^{-1} due east. This is illustrated in figure 9.5b.

SAMPLE PROBLEM 9.4

Velocity calculation

A cyclist is riding along a straight road at a constant velocity of 36 km h^{-1} (10 m s^{-1}) in an easterly direction. A car approaches the cyclist from behind and is initially 360 m behind the cyclist. If the car is travelling at a speed of 100 km h^{-1} (28 m s^{-1}), how long will it take to catch up to the cyclist?

SOLUTION

The velocity of the car relative to the cyclist is the difference between the velocity of the car relative to the ground and the velocity of the cyclist relative to the ground. That is, 28 m s^{-1} due east minus 10 m s^{-1} due east equals 18 m s^{-1} due east. The time taken can be calculated using the formula:

$$v_{av} = \frac{\Delta \boldsymbol{r}}{\Delta t}.$$

$$18 \text{ m s}^{-1} \text{ due east} = \frac{360 \text{ m due east}}{t}$$

$$t = \frac{360}{18}$$

$$= 20 \text{ s}.$$

PHYSICS FACT

Do you always feel like you're on the move? No wonder! When you are standing still, you are actually moving through space at a speed of about 30 km s^{-1}. That's about 110 000 km h^{-1}! This is the speed at which the Earth is hurtling through space in orbit around the Sun. An observer on the Sun could measure that speed. If you were standing still in Sydney, a person high above the South Pole would say that you were rotating with the ground around the Earth's axis at a speed of over 1300 km h^{-1}.

The speed you measure depends on your position, how fast you are moving and your direction of movement.

Instantaneous speed and velocity

Neither the average speed nor the average velocity provides information about movement at any particular instant of time. For example, if you were to drive from Sydney to Dubbo, a distance of 400 km, in 5 hours, your average speed would be 80 km h^{-1}. However, that doesn't mean that you have travelled at a constant speed of 80 km h^{-1} for the entire trip.

Instantaneous speed is the speed at a particular instant of time.

Instantaneous velocity is the velocity at a particular instant of time.

The speed at any particular instant of time is called the **instantaneous speed**. The velocity at any particular instant of time is, not surprisingly, called the **instantaneous velocity**. If an object moves with a constant velocity during a time interval, its instantaneous velocity throughout the interval is the same as its average velocity.

PHYSICS IN FOCUS

As we saw in chapter 3, page 41, radar guns and mobile radar units in police cars detect speeding cars by measuring instantaneous speed. This method of measuring speed doesn't require a measure of distance or time. Radio waves emitted from a radar gun or unit are reflected from the target vehicle. The reflected waves have a different frequency from the emitted waves because of the movement of the vehicle. The faster the vehicle is travelling, the more the frequency changes. The altered waves are detected by a receiver. Some radar units are connected to speed cameras that automatically photograph speeding vehicles.

Laser guns measure average speed. They emit light pulses which are reflected from the target vehicle. The time taken for each pulse to return is recorded and compared with that of the previous pulse. This allows the average speed over a very small time interval to be calculated. Laser guns are useful in heavy traffic because the beam is much narrower than the 'beam' of radio waves emitted by radar guns and units. That means that individual vehicles suspected of speeding can be targeted.

Figure 9.6
Police use laser guns to measure the average speed of vehicles.

9.3 ACCELERATION

Acceleration is the time rate of change of velocity.

The rate at which an object changes its velocity is called its **acceleration**. Because velocity is a vector quantity, it follows that acceleration is also a vector quantity. The direction of the acceleration of an object is the same as the direction of its change in velocity.

The average acceleration of an object, $\boldsymbol{a}_{av}$, can be expressed as:

$$\boldsymbol{a}_{av} = \frac{\Delta \boldsymbol{v}}{\Delta t}$$

where

$\Delta \boldsymbol{v}$ = the change in velocity during the time interval Δt.

The change in velocity is found by subtracting the initial velocity, $\boldsymbol{u}$, from the final velocity, $\boldsymbol{v}$.

Thus: $$\boldsymbol{a}_{av} = \frac{\boldsymbol{v} - \boldsymbol{u}}{t}$$

where

t is the time during which the change in velocity occurs.

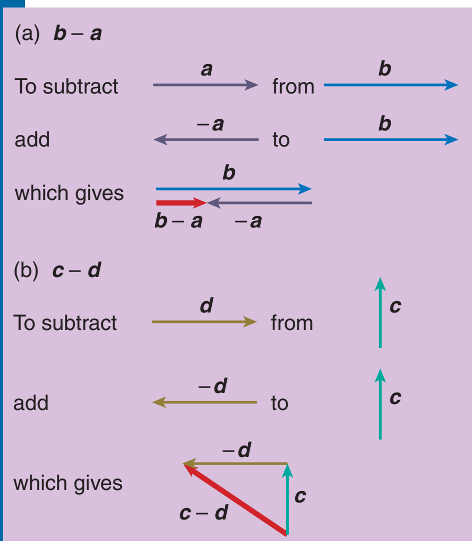


Figure 9.7 *Subtracting vectors*

Subtracting vectors

In order to determine a change in velocity it is necessary to subtract the vector $\boldsymbol{u}$ from the vector $\boldsymbol{v}$. One vector can be subtracted from another by simply adding its negative. This works because subtracting a vector is the same as adding the negative vector (just as subtracting a positive number is the same as adding the negative of that number). The method of adding vectors is shown on page 173. Two examples of vector subtraction are shown in figure 9.7.

SAMPLE PROBLEM 9.5

Acceleration of a car

A car starts from rest and reaches a velocity of 20 m s^{-1} due east in 5.0 s. What is its average acceleration?

SOLUTION

$$\boldsymbol{a}_{av} = \frac{\boldsymbol{v} - \boldsymbol{u}}{t}$$

$\boldsymbol{v} - \boldsymbol{u}$ = 20 m s^{-1} due east – 0
= 20 m s^{-1} due east (no vector diagram needed here)

$$\boldsymbol{a}_{av} = \frac{20 \text{ m s}^{-1} \text{ due east}}{5.0 \text{ s}}$$
= 4.0 m s^{-2} due east

SAMPLE PROBLEM 9.6

Acceleration of a cyclist

What is the average acceleration of a cyclist riding north who slows down from a speed of 8.0 m s^{-1} to a speed of 5.0 m s^{-1} in 2.0 s?

SOLUTION

$$\boldsymbol{a}_{av} = \frac{\boldsymbol{v} - \boldsymbol{u}}{t}$$

$\boldsymbol{v} - \boldsymbol{u}$ = 5.0 m s^{-1} north – 8.0 m s^{-1} north
= –3.0 m s^{-1} north (no vector diagram needed here)

$$\boldsymbol{a}_{av} = \frac{-3.0 \text{ m s}^{-1} \text{ north}}{2.0 \text{ s}}$$
= –1.5 m s^{-2} north

A negative acceleration is called a deceleration. This acceleration could also be expressed as 1.5 m s^{-2} south.

SAMPLE PROBLEM 9.7

Acceleration of a bus

What is the average acceleration of a bus that travels west at 15 m s^{-1} and then turns right to drive north at a speed of 20 m s^{-1}? The turn is completed in a time interval of 2.5 s.

SOLUTION

The change in velocity must first be found by subtracting vectors since:

$\Delta \boldsymbol{v} = \boldsymbol{v} - \boldsymbol{u}$.

The magnitude of the change in velocity can be found by using Pythagoras' theorem or by trigonometry.

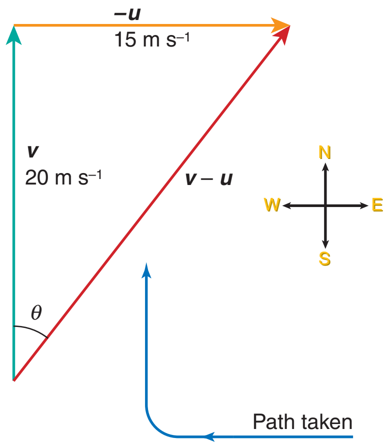


Figure 9.8

$$\Delta v = \sqrt{(20 \text{ m s}^{-1})^2 + (15 \text{ m s}^{-1})^2}$$
$$= 25 \text{ m s}^{-1}$$

The direction can be found by calculating the value of the angle θ.

$$\tan \theta = \frac{15}{20} = 0.75$$
$$\theta = 37°$$

The direction of the change in velocity is therefore N37°E.
The average acceleration of the bus is given by:

$$\boldsymbol{a}_{av} = \frac{\boldsymbol{v} - \boldsymbol{u}}{\Delta t}$$
$$= \frac{25 \text{ m s}^{-1} \text{ N37°E}}{2.5 \text{ s}}$$
$$= 10 \text{ m s}^{-2} \text{ N37°E.}$$

eModelling:
Numerical model for acceleration
doc-0050

PHYSICS FACT

In drag racing, cars are able to reach speeds of about 420 km h^{-1} from a standing start over a distance of 400 m. They can cover the distance in less than 6 s. This represents an average acceleration of more than 70 km h^{-1} s^{-1} or about 20 m s^{-2}. The initial acceleration can be about three times larger than this.

The acceleration of a world-class athlete at the beginning of a 100 m sprint is about 3 m s^{-2}. The fastest land animal, the cheetah, takes 2 s to reach its maximum speed of about 30 m s^{-1}. What is its average acceleration during this period?

Figure 9.9 *The powerful leg and shoulder muscles of the cheetah give it extraordinary acceleration.*

eBookplus

eLesson:
Motion with constant acceleration
eles-0030

Constant acceleration formulae

When acceleration is constant (including when it is zero), the motion of an object can be described by some simple formulae. The definition of average acceleration leads to the first of these formulae. When the acceleration is constant, its value is the same as the average acceleration:

$$\boldsymbol{a} = \boldsymbol{a}_{av} = \frac{\Delta \boldsymbol{v}}{\Delta t}$$

where
$\Delta \boldsymbol{v}$ = the change in velocity during the time interval Δt.
When $t = 0$ the velocity is $\boldsymbol{u}$.

Thus:
$$\boldsymbol{a} = \frac{\boldsymbol{v} - \boldsymbol{u}}{t}$$

where
$\boldsymbol{v}$ is the velocity at time t.

$$\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{a}t$$
$$\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{a}t$$

Note that this equation is a vector equation. The direction of the change in velocity ($\boldsymbol{v} - \boldsymbol{u}$) is the same as the direction of the acceleration. As long as the motion is along a straight line, the vectors can be expressed as positive or negative quantities. Vector notation is not necessary.

Thus:
$$v = u + at. \qquad [1]$$

The second of the constant acceleration equations can be found by restating the definition of average velocity.

$$\boldsymbol{v}_{av} = \frac{\Delta \boldsymbol{r}}{\Delta t} = \frac{\boldsymbol{r}}{t}$$

where
$\boldsymbol{r}$ = displacement from starting position at time t.
When the acceleration is constant, the average velocity can be expressed as $\boldsymbol{v}_{av} = \dfrac{\boldsymbol{u} + \boldsymbol{v}}{2}$.

Thus:

$$\frac{\boldsymbol{u} + \boldsymbol{v}}{2} = \frac{\boldsymbol{r}}{t}$$

$$r = \tfrac{1}{2}(u + v)\,t \qquad [2]$$

Once again, vector notation is not necessary as long as the motion is along a straight line.

A third formula can be obtained by combining formulae [1] and [2]. Substituting v from formula [1] into formula [2] gives

$$r = \tfrac{1}{2}(u + u + at)\,t$$

$$= \tfrac{1}{2}(2u + at)\,t$$

$$r = ut + \tfrac{1}{2}at^2 \qquad [3]$$

SAMPLE PROBLEM 9.8

Velocity and distance calculations

A physics student drops a coin into a wishing well and takes 3.0 s to make a wish (for a perfect score in the next physics test!). The coin splashes into the water just as she finishes making her wish. The coin accelerates towards the water at a constant 9.8 m s^{-2}.

(a) What is the coin's velocity as it strikes the water?
(b) How far does the coin fall before hitting the water?

SOLUTION

(a) $u = 0$

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

$t = 3.0 \text{ s}$

The appropriate formula here is $v = u + at$.

$$v = 0 + 9.8 \times 3.0$$
$$= 29.4 \text{ m s}^{-1}$$

The coin is travelling at a velocity of 29 m s^{-1} down as it strikes the water.

(b) The appropriate formula here is $r = ut + \frac{1}{2}at^2$ because it includes the three known quantities along with the unknown quantity r.

$$r = 0 + \tfrac{1}{2} \times 9.8 \times (3.0)^2$$
$$= 44.1 \text{ m}$$

The coin falls a distance of 44 m.

SAMPLE PROBLEM 9.9

Speed and acceleration of a skidding car

The driver of a car travelling along a suburban street was forced to brake suddenly to prevent serious injury to the neighbour's cat. The car skidded in a straight line for 2.0 s, stopping just a millimetre or two away from the cat. The deceleration was constant and the length of the skid mark was 12 m.

(a) At what speed was the car travelling as it began to skid?
(b) What was the acceleration of the car?

SOLUTION

(a) $r = 12$ m, $t = 2.0$ s, $v = 0$ (assigning forward as positive)

The appropriate formula is:

$$r = \tfrac{1}{2}(u + v)\ t$$
$$12 = \tfrac{1}{2}(u + 0)\ 2.0$$
$$u = 12 \text{ m s}^{-1}.$$

The car was travelling at a speed of 12 m s^{-1}. That's about 43 km h^{-1}.

(b) The appropriate formula is:

$$v = u + at$$
$$0 = 12 + a \times 2.0$$
$$a = \frac{-12}{2.0}$$
$$= -6.0 \text{ m s}^{-2}.$$

The acceleration of the car was -6.0 m s^{-2}.

9.4 GRAPHING MOTION

A description of motion in terms of displacement, average velocity and average acceleration is not really complete. These quantities provide a 'summary' of motion, but do not usually provide detailed information about the velocity and acceleration at any particular instant of time. By describing the motion of an object in a graphical form, it is possible to estimate the displacement, velocity and acceleration of an object or person at any instant during a chosen time interval.

Graphing displacement versus time

'Bolter' Beryl and 'Steady' Sam decided to race over a distance of 100 metres. They run due west. Both runners were already moving before the start of the first interval but they did cross the starting line at the same time. Timekeepers were instructed to record the position of each runner after each 3.0 second interval, as shown in table 9.1. A graph of displacement versus time, representing the motion of both runners, was then plotted (see figure 9.10).

Table 9.1 **The progress of Beryl and Sam**

	DISPLACEMENT (MEASURED FROM STARTING LINE) (m)	
TIME (s)	'BOLTER' BERYL	'STEADY' SAM
0.0	0	0
3.0	43	20
6.0	64	40
9.0	78	60
12.0	90	80
15.0	100	100

A number of observations can be made from the graph of displacement versus time:

- Both runners reach the finish at the same time. The result is a dead heat. Beryl and Sam have the same average speed and the same average velocity.
- 'Steady' Sam manages to maintain a constant velocity throughout the race. In fact his instantaneous velocity at every instant throughout the race is the same as his average velocity. Sam's average velocity and instantaneous velocity are both equal to the gradient of the displacement–time graph since:

$$\boldsymbol{v}_{\text{av}} = \frac{\Delta \boldsymbol{r}}{\Delta t}$$

$$= \frac{100 \text{ m west}}{15 \text{ s}} = \frac{\text{rise}}{\text{run}} = \text{gradient.}$$

  Sam's average velocity is 6.7 m s^{-1} west. His instantaneous velocity at every instant throughout the race is also 6.7 m s^{-1} west.
- 'Bolter' Beryl makes a flying start but, after her initial 'burst', her instantaneous velocity decreases throughout the race as she tires. Beryl's average velocity cannot be determined by calculating the gradient because the gradient changes. Beryl's average velocity for the whole race can be determined using the formula $\boldsymbol{v}_{\text{av}} = \frac{\Delta \boldsymbol{r}}{\Delta t}$. Her average velocity is therefore 6.7 m s^{-1} west, the same as Sam's. However, Beryl's instantaneous velocity changes throughout the race.

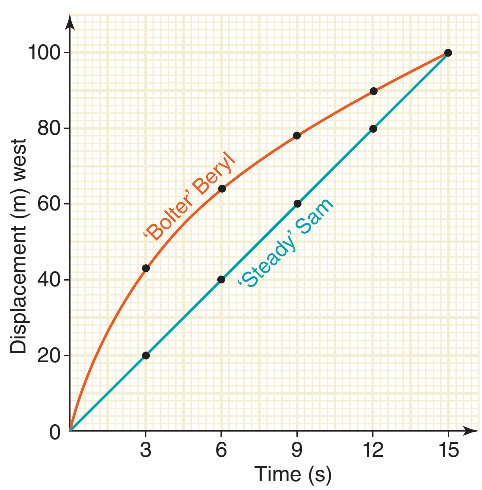


Figure 9.10 ***The graph of displacement versus time provides valuable information about the way that the race was run.***

A more detailed description of Beryl's motion can be given by calculating her average velocity during each 3.0 second interval of the race. The results are shown in table 9.2.

Table 9.2 Beryl's changing velocity

TIME INTERVAL (s)	DISPLACEMENT Δr (m west)	AVERAGE VELOCITY DURING INTERVAL $v_{av} = \frac{\Delta r}{\Delta t}$ (m s^{-1} west)
0–3.0	$43 - 0 = 43$	14
3.0–6.0	$64 - 43 = 21$	7.0
6.0–9.0	$78 - 64 = 14$	4.7
9.0–12.0	$90 - 78 = 12$	4.0
12.0–15.0	$100 - 90 = 10$	3.3

The average velocity during each interval is the same as the gradient of the straight line joining the data points representing the beginning and end of the interval.

Figure 9.11 shows how small time intervals can be used to find Beryl's instantaneous velocity at an instant 4.0 seconds from the start of the race. Her instantaneous velocity is not the same as the average velocity during the 3.0–6.0 second time interval shown in table 9.2. However, it can be estimated by drawing the line AD and finding its gradient. The gradient of the line BC would provide an even better estimate of the instantaneous velocity. If you continue this process of decreasing the time interval used to estimate the instantaneous velocity, you will eventually obtain a line that is a tangent to the curve. The gradient of the tangent is the same as the gradient of the point on the curve at which it is drawn. It is therefore equal to the instantaneous velocity at the instant represented by the point on the curve. The gradient of the tangent to the curve at 4.0 seconds in figure 9.11 can be determined by using the points P and Q.

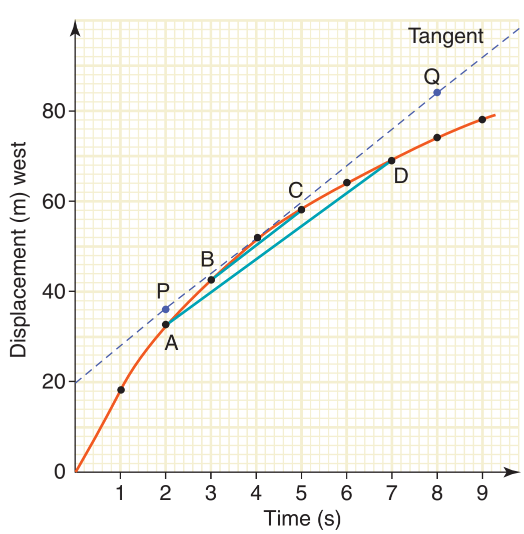


Figure 9.11 *The first 9.0 seconds of 'Bolter' Beryl's run*

That is, gradient $= \frac{\text{rise}}{\text{run}}$

$$= \frac{(84 - 36)\ \text{m}}{(8.0 - 2.0)\ \text{s}}$$

$$= \frac{48}{6.0}$$

$$= 8.0\ \text{m s}^{-1}.$$

The instantaneous velocity of an object can be obtained from the gradient of its displacement-versus-time graph.

'Bolter' Beryl's instantaneous velocity at 4.0 seconds from the start of the race is therefore 8.0 m s^{-1} west.

Just as the gradient of a position-versus-time graph can be used to determine the velocity of an object, the gradient of a distance-versus-time graph can be used to determine its speed. Because Beryl and Sam were running in a straight line and in one direction only, their distance from the starting point is the magnitude of their displacement. Their speed is equal to the magnitude of their velocity.

SAMPLE PROBLEM 9.10

Straight-line motion of a cyclist

The graph in figure 9.12 is a record of the straight-line motion of a cyclist who initially moves in a northerly direction. Use the graph to determine:

(a) the displacement of the cyclist at the end of the 60 s interval

(b) the distance travelled by the cyclist during the 60 s interval

(c) the average velocity of the cyclist over the first 20 s

(d) the instantaneous velocity of the cyclist 20 s from the start of the motion.

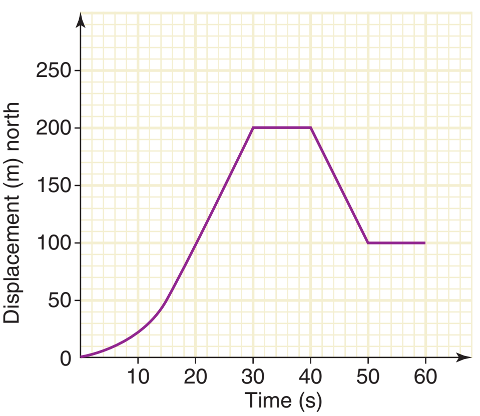


Figure 9.12 *The straight-line motion of a cyclist*

SOLUTION

(a) The displacement can be read directly from the graph. After 60 s it is 100 m north.

(b) The distance travelled by the cyclist = 200 + 100
= 300 m.

(c) $$\boldsymbol{v}_{\text{av}} = \frac{\Delta \boldsymbol{r}}{\Delta t}$$

$$= \frac{100\ \text{m north}}{20\ \text{s}}$$

$$= 5.0\ \text{m s}^{-1}\ \text{north}$$

(d) The instantaneous velocity is equal to the gradient of the displacement-versus-time graph. Twenty seconds from the start of the motion the gradient is equal to $\frac{\text{rise}}{\text{run}}$:

$$= \frac{200 - 50}{30 - 15}$$

$$= \frac{150}{15}$$

$$= 10\ \text{m s}^{-1}\ \text{north}.$$

Graphing velocity versus time

The race between 'Bolter' Beryl and 'Steady' Sam can also be described by a graph of velocity versus time. Sam's velocity is 6.7 m s^{-1} due west throughout the race. The curve describing Beryl's motion can be plotted by determining the instantaneous velocity at various times during the race. This can be done by drawing tangents at a number of points on the displacement-versus-time graph in figure 9.10 (page 182). Table 9.3 shows the data obtained using this method. The velocity-versus-time graph in figure 9.13 describes the motion of Beryl and Sam.

Table 9.3 Beryl's velocity during the race

TIME (s)	VELOCITY (m s^{-1} west) (OBTAINED FROM THE GRADIENT OF HER DISPLACEMENT-VERSUS-TIME GRAPH)
0.0	18
2.0	12
4.0	8.0
6.0	5.4
8.0	4.7
10.0	4.2
12.0	3.5
14.0	3.1

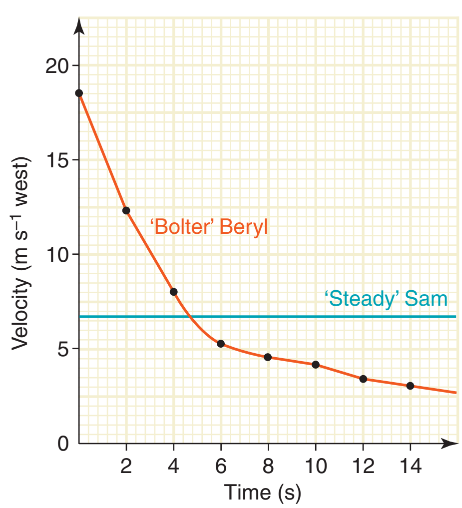


Figure 9.13 *Graph of velocity versus time for the race*

The velocity-versus-time graph shows that Sam's velocity is constant and equal to his average velocity. It allows you to estimate the velocity of each runner at any time. It also confirms that both runners are actually moving when they cross the starting line. It provides a much clearer picture of the way that Beryl's velocity changes during the race. Note that:

- The magnitude of her velocity decreases rapidly at first but less rapidly towards the end of the race.
- For most of the duration of the race, she is running more slowly than Sam. In fact Beryl's speed (the magnitude of her velocity) drops below Sam's after only 4.7 seconds.

Displacement from a velocity-versus-time graph

In the absence of a position–time graph, a velocity-versus-time graph provides useful information about the change in position, or displacement of an object. Sam's constant velocity, the same as his average velocity, makes it very easy to determine his displacement during the race.

$$\begin{aligned}\Delta \boldsymbol{r} &= \boldsymbol{v}_{av}\Delta t \qquad \left(\text{since } \boldsymbol{v}_{av} = \frac{\Delta \boldsymbol{r}}{\Delta t}\right)\\ &= 6.7\text{ m s}^{-1}\text{ west} \times 15\text{ s}\\ &= 100\text{ m west}\end{aligned}$$

This displacement is equal to the area of the rectangle under the graph describing his motion.

$$\begin{aligned}\text{Area} &= \text{length} \times \text{width} = 15\text{ s} \times 6.7\text{ m s}^{-1}\text{ west}\\ &= 100\text{ m west}\end{aligned}$$

Because the race was a dead heat, we know that Beryl's average velocity was also 6.7 m s^{-1}. Her displacement during the race can be calculated in the same way as Sam's.

$$\begin{aligned}\Delta \boldsymbol{r} &= \boldsymbol{v}_{av}\Delta t\\ &= 6.7\text{ m s}^{-1} \times 15\text{ s}\\ &= 100\text{ m west}\end{aligned}$$

The displacement of an object during a time interval can be obtained by determining the area under the velocity-versus-time graph for that time interval.

Beryl's displacement can also be found by calculating the area under the velocity-versus-time graph that describes her motion. This can be done by 'counting squares' or by dividing the area under the graph into rectangles and triangles. The area under Beryl's velocity-versus-time graph is, not surprisingly, 100 m. In fact, the area under any part of the velocity-versus-time graph is equal to the displacement during the interval represented by that part.

SAMPLE PROBLEM 9.11

Using speed to determine distance travelled

In the race between 'Bolter' Beryl and 'Steady' Sam, how far ahead of Sam was Beryl when her speed dropped below Sam's speed?

SOLUTION

Although it is possible to answer this question using the displacement-versus-time graph in figure 9.10 (you might like to explain how you would do this), it can also be done using the velocity-versus-time graph in figure 9.13. It shows that Beryl's speed (and magnitude of her velocity) drops below Sam's 4.7 seconds after the race starts.

Sam's displacement after 4.7 s is equal to the area under the line representing the first 4.7 s of his motion. That is, 4.7 s × 6.7 m s^{-1} west. Therefore, Sam is 31 m west of the starting line after 4.7 s.

Beryl's displacement after 4.7 s is equal to the area under the curve representing the first 4.7 s of her motion. This area can be estimated by determining the shaded area P and rectangle Q shown in figure 9.14. The area P can be assumed to be a triangle for this purpose.

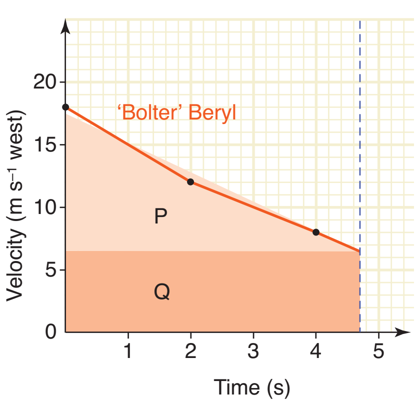


Figure 9.14 *Calculating displacement using a velocity-versus-time graph*

$$\begin{aligned}\text{Area} &= \text{area P} + \text{area Q}\\ &= \tfrac{1}{2} \times 4.7 \times 11 + 4.7 \times 6.7\\ &= 25.85 + 30.55\\ &= 56.40\text{ m west}\end{aligned}$$

Beryl is therefore 56 m west of the starting line after 4.7 s. She is 56 − 31 = 25 m ahead of Sam when her speed drops below his.

Acceleration from a velocity-versus-time graph

PRACTICAL ACTIVITIES

9.2

On your bike

The graph in figure 9.15 describes the motion of an elevator as it moves from the ground floor to the top floor and back down again. It stops briefly at the top floor to pick up a passenger. For convenience, any upward displacement from the ground floor is defined as positive. The graph has been divided into seven sections labelled A–G.

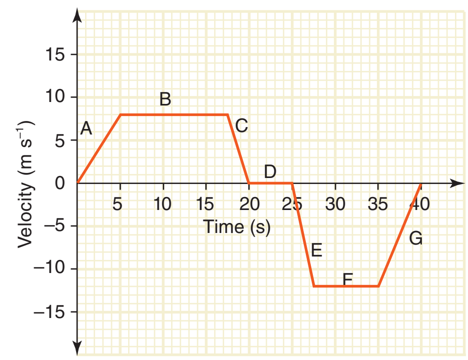


Figure 9.15 *The motion of an elevator*

The acceleration at any instant during the motion can be determined by calculating the gradient of the graph. This is a consequence of the definition of acceleration. When the acceleration is constant:

The acceleration of an object can be obtained from the gradient of its velocity-versus-time graph.

$$\boldsymbol{a} = \boldsymbol{a}_{av} = \frac{\Delta \boldsymbol{v}}{\Delta t}$$

$$= \text{gradient of a velocity-versus-time graph.}$$

Throughout interval A, the acceleration, $\boldsymbol{a}$, of the elevator is:

$$\boldsymbol{a} = \frac{\text{rise}}{\text{run}} = \frac{+8.0}{5.0}$$
$$= +1.6 \text{ m s}^{-2} \text{ or } 1.6 \text{ m s}^{-2} \text{ up.}$$

PRACTICAL ACTIVITIES

9.3

Analysing motion with a constant acceleration

During intervals B, D and F the velocity is constant and the gradient of the graph is zero. The acceleration during each of these intervals is, therefore, zero. Throughout interval C, the acceleration is:

$$\boldsymbol{a} = \frac{-8.0}{2.5} = -3.2 \text{ m s}^{-2} \text{ or } 3.2 \text{ m s}^{-2} \text{ down.}$$

Throughout interval E, the acceleration is:

$$\boldsymbol{a} = \frac{-12}{2.5} = -4.8 \text{ m s}^{-2} \text{ or } 4.8 \text{ m s}^{-2} \text{ down.}$$

***SAMPLE PROBLEM* 9.12**

Using a velocity-versus-time graph to determine acceleration and displacement

Use the velocity-versus-time graph in figure 9.15 to answer the following questions.

(a) What was the acceleration of the elevator during interval G?
(b) When did the elevator first change direction?
(c) What was the displacement of the elevator after 40 s?

SOLUTION

(a) $\boldsymbol{a} = \text{gradient}$

$$= \frac{+12}{5.0}$$
$$= +2.4 \text{ m s}^{-2} \text{ or } 2.4 \text{ m s}^{-2} \text{ up}$$

(b) The initial velocity is positive, which corresponds with the elevator moving up. The first movement downwards occurs at the end of interval D, when the velocity first becomes negative. During interval C, even though the acceleration is negative, the velocity is still positive, but decreasing.

(c) The displacement of the elevator after 40 s is equal to the area under the graph in figure 9.15. Dividing the area under the graph into triangles and rectangles and working from left to right, yields an area of:

$$\left(\tfrac{1}{2} \times 5.0 \times 8.0\right) + (12.5 \times 8.0) + \left(\tfrac{1}{2} \times 2.5 \times 8.0\right) + \left(\tfrac{1}{2} \times 2.5 \times -12\right) + (7.5 \times -12) + \left(\tfrac{1}{2} \times 5.0 \times -12\right)$$
$$= 20 + 100 + 10 - 15 - 90 - 30$$
$$= -5.0 \text{ m.}$$

This represents a downwards displacement of 5.0 m, which is consistent with the elevator finally stopping two floors below the ground floor.

PHYSICS FACT

A non-zero acceleration does not always result from a change in speed. Consider a car travelling at 60 km h^{-1} in a northerly direction turing right and continuing in an easterly direction at the same speed. Assume that the complete turn takes 10 s. The average acceleration during the time interval of 10 s is given by:

$$\boldsymbol{a}_{\text{av}} = \frac{\Delta \boldsymbol{v}}{\Delta t}.$$

The change in velocity must be determined first. Thus,

$$\Delta \boldsymbol{v} = \boldsymbol{v} - \boldsymbol{u}$$
$$= \boldsymbol{v} + -\boldsymbol{u}.$$

The vectors $\boldsymbol{v}$ and $-\boldsymbol{u}$ are added together to give the resulting change in velocity.

The magnitude of the change in velocity is calculated using Pythagoras' Theorem or trigonometric ratios to be 85 km h^{-1}. Alternatively, the vectors can be added using a scale drawing and then measuring the magnitude and direction of the sum. The direction of the change in velocity can be seen in figure 9.16 to be south-east.

$$\boldsymbol{a}_{\text{av}} = \frac{\Delta \boldsymbol{v}}{\Delta t}$$
$$= \frac{85}{10}$$
$$= 8.5 \text{ km h}^{-1} \text{ s}^{-1} \text{ south-east}$$

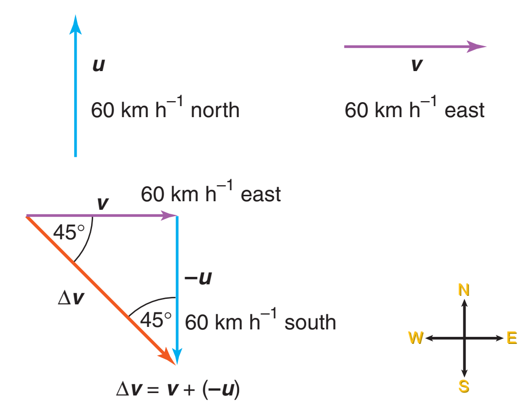


Figure 9.16 *A change in velocity can occur even if there is no change in speed.*

SUMMARY

- Displacement is a measure of the change in position of an object. Displacement is a vector quantity.
- In order to fully describe any vector quantity, a direction must be specified as well as a magnitude.
- Speed is a measure of the rate at which an object moves over distance and is a scalar quantity. Velocity is the time rate of displacement and is a vector quantity.
- The velocity of an object measured by a moving observer is referred to as the relative velocity. The relative velocity is the difference between the velocity of the object relative to the ground and the velocity of the observer relative to the ground.
- Average speed = $\dfrac{\text{distance travelled}}{\text{time interval}}$.
- Average velocity = $\dfrac{\text{displacement}}{\text{time interval}}$. The average velocity of an object, $\boldsymbol{v}_{\text{av}}$ during a time interval, t, can be expressed as $\boldsymbol{v}_{\text{av}} = \dfrac{\Delta \boldsymbol{r}}{\Delta t}$.
- Instantaneous speed is the speed at a particular instant of time. Instantaneous velocity is the velocity at a particular instant of time.
- Acceleration is the rate at which an object changes its velocity. Acceleration is a vector quantity. The average acceleration of an object, $\boldsymbol{a}_{\text{av}}$, can be expressed as $\boldsymbol{a}_{\text{av}} = \dfrac{\Delta \boldsymbol{v}}{\Delta t}$, where $\Delta \boldsymbol{v}$ = the change in velocity during the time interval Δt.
- When the acceleration of an object is constant, the following formulae can be used to describe its motion:
  $v = u + at$
  $r = \frac{1}{2}(u + v)t$
  $r = ut + \frac{1}{2}at^2$.
- The instantaneous velocity of an object can be found from a graph of its displacement versus time by calculating the gradient of the graph. Similarly, the instantaneous speed can be found from a graph of its distance versus time by calculating the gradient of the graph.
- The displacement of an object during a time interval can be found by determining the area under its velocity-versus-time graph. Similarly, the distance travelled by an object can be found by determining the area under its speed-versus-time graph.
- The instantaneous acceleration of an object can be found from a graph of its velocity versus time by calculating the gradient of the graph.

QUESTIONS

1. State which of the following are vector quantities.
   (a) distance
   (b) displacement
   (c) speed
   (d) velocity
2. After a car is driven 5 km in a southerly direction it turns right and travels 12 km in a westerly direction. The journey takes 12 min.
   (a) Calculate the distance it has travelled.
   (b) Calculate its displacement.
   (c) Calculate the car's average speed.
   (d) Calculate the car's average velocity.
3. In 2004, cyclist Sarah Ulmer, of New Zealand, set a world record of 3 min 24.537 s for the 3000 m pursuit.
   (a) Calculate her average speed.
   (b) Calculate the time taken by Sarah to cycle from Sydney to Newcastle, a distance of 155 km, if she could maintain her average speed for the 3000 m pursuit for the whole distance.
   (c) Calculate the time for a car to travel from Sydney to Newcastle if its average speed is 80 km h^{-1}.
   (d) A car travels from Sydney to Newcastle and back to Sydney in 4.0 h.
      (i) Calculate its average speed.
      (ii) Calculate its average velocity.
4. Explain why you cannot ever measure the instantaneous velocity of an object with a stopwatch.
5. A holidaying physics teacher drives her old Volkswagen from Sydney to Batemans Bay, a distance of 300 km. Her average speed is 80 km h^{-1}. She trades in her Volkswagen and purchases a brand new Toyota Corolla. She proudly drives her new car back home to Sydney at an average speed of 100 km h^{-1}.

(a) Make a quick prediction of her average speed for the whole trip.

(b) Calculate the average speed for the whole journey and explain any difference between the predicted and calculated average speed.

6. Which is larger in magnitude — the speed of a fly or the velocity of a fly? Explain your answer.

7. The police are pursuing a speeding motorist on a straight road. The speeding car is travelling at 90 km h^{-1} (25 m s^{-1}). The police car, initially 200 metres behind the speeding car, travels at a speed of 105 km h^{-1} (29 m s^{-1}) with lights flashing and siren screaming. Calculate how long it takes the police car to catch up with the speeding car.

8. Calculate the time for:

(a) a car to accelerate on a straight road at a constant 6.0 m s^{-2} from an initial speed of 60 km h^{-1} (17 m s^{-1}) to a final speed of 100 km h^{-1} (28 m s^{-1})

(b) a cyclist to accelerate from rest at a constant 2.0 m s^{-2} to a speed of 10 m s^{-1}.

9. Calculate (i) the change in speed and (ii) the change in velocity in each of the following situations.

(a) The driver of a car heading north along a freeway at 100 km h^{-1} slows down to 60 km h^{-1} as the traffic gets heavier.

(b) A fielder catches a cricket ball travelling towards him at 20 m s^{-1}.

(c) A tennis ball travelling at 25 m s^{-1} is returned directly back to the server at a speed of 30 m s^{-1}.

10. A car travelling east at a speed of 100 m s^{-1} turns right to head south at the same speed. Has the car undergone an acceleration? Explain your answer with the aid of a diagram.

11. Calculate the average acceleration of:

(a) a car, starting from rest, which reaches a velocity of 20 m s^{-1} due north in 5.0 s

(b) a cyclist travelling due west at a speed of 15 m s^{-1}, who turns to cycle due north at a speed of 20 m s^{-1} (the change occurs in a time interval of 2.5 s)

(c) a bus travelling due north at 8.0 m s^{-1}, which turns right to travel due east without changing speed, in a time interval of 4.0 s.

12. In Acapulco, on the coast of Mexico, professional high divers plunge from a height of 36 m above the water. (The highest diving boards in Olympic diving events are 10 m above the water.) Estimate:

(a) the length of the time interval during which the divers fall through the air

(b) the speed with which the divers enter the water.

Assume that throughout their dive, the divers are falling vertically from rest with an acceleration of 9.8 m s^{-2}.

13. A skateboard rider travelling down a hill notices the busy road ahead and comes to a stop in 2.0 s over a distance of 12 m. Assume a constant negative acceleration.

(a) Calculate the initial speed of the skateboard.

(b) Calculate the acceleration of the skateboard as it came to a stop.

14. A car is travelling at a speed of 100 km h^{-1} (27.8 m s^{-1}) when the driver sees a large fallen tree branch in front of her. At the instant that she sees the branch it is 50.0 m from the front of her car. After she applies the brakes, the car travels a distance of 48.0 m before coming to a stop.

(a) Calculate the time taken for the car to stop once the brakes were applied.

(b) Calculate the average acceleration of the car while it is braking.

(c) What other information do you need in order to determine whether the car stops before it hits the tree branch? Make an estimate of the missing item of information to predict whether or not the car is able to stop in time.

15. Amy rides a toboggan down a steep snow-covered slope. Starting from rest, Amy reaches a speed of 12 m s^{-1} as she passes her brother, who is standing 19 m further down the slope from her starting position. Assume that Amy's acceleration is constant.

(a) Calculate the time taken for Amy to reach her brother.

(b) Calculate Amy's acceleration.

(c) At what instant was Amy's instantaneous velocity equal to her average velocity?

16. The position-versus-time graph in figure 9.17 describes the motion of six different objects labelled A–E.

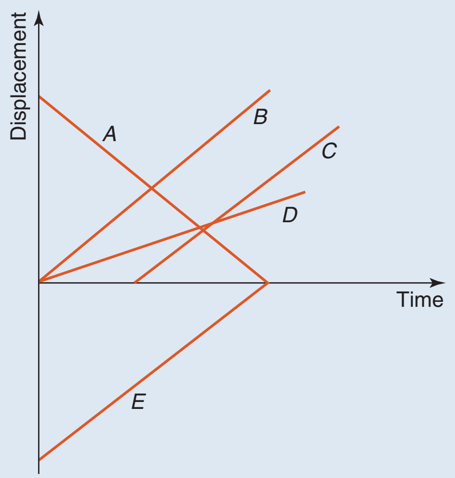


Figure 9.17

(a) Which two objects start from the same position, but at different times?

(b) Which two objects start at the same position at the same time?

(c) Which two objects are travelling at the same speed as each other, but with different velocities?

(d) Which two objects are moving towards each other?

(e) Which object has a lower speed than all of the other objects?

17. The velocity-versus-time graph in figure 9.18 describes the motion of a car as it travels due south through an intersection. The car was stationary for 6 s while the traffic lights were red.

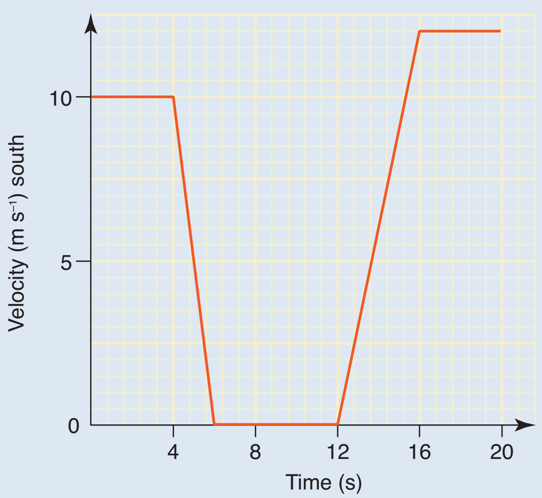


Figure 9.18 ***The straight-line motion of a car travelling through an intersection***

(a) Calculate the displacement of the car during the time interval in which it was slowing down.

(b) Calculate the average acceleration of the car during the time interval in which it was slowing down.

(c) Calculate the average acceleration of the car during the first 4.0 s after the lights turned green.

(d) Calculate the average velocity of the car during the time interval described by the graph.

18. The graph in figure 9.19 is a record of the straight-line motion of a skateboard rider during an 80 s time interval. The interval has been divided into sections A–E. The skateboarder initially moves north from the starting point.

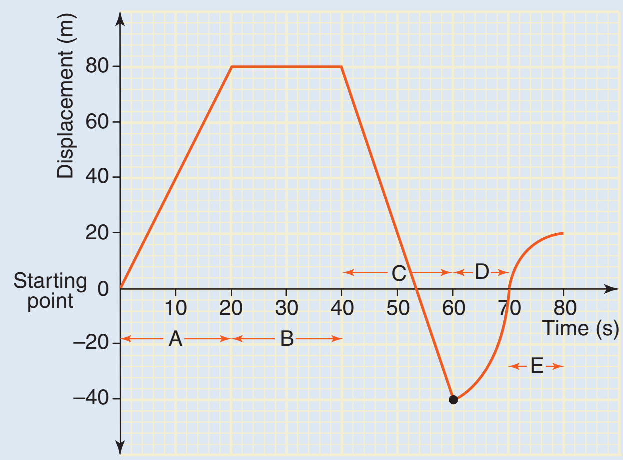


Figure 9.19

(a) During which section of the interval was the skateboard rider stationary?

(b) During which sections was the skateboarder travelling north?

(c) At what instant did the skateboard rider first move back towards the starting line?

(d) What was the total displacement of the skateboarder after the 80 s interval?

(e) What distance did the skateboarder travel during the 80 s interval?

(f) During which section was the skateboard rider speeding up?

(g) During which section was the skateboard rider slowing down?

(h) What was the skateboarder's average speed during the entire interval?

(i) What was the velocity of the skateboarder throughout section C?

(j) Estimate the velocity of the skateboarder 65 s into the interval.

19. Sketch a velocity-versus-time graph to illustrate the motion described in each of the following situations.
   (a) A bicycle is pedalled steadily along a road. The cyclist stops pedalling and allows the bicycle to come to a stop.
   (b) A ball is thrown straight up into the air and is caught at the same height from which it was thrown. The acceleration of the ball is constant and downwards.

20. The graph in figure 9.20 is a record of the motion of a remote-controlled car during an 80 s time interval. The interval has been divided into sections A–G.

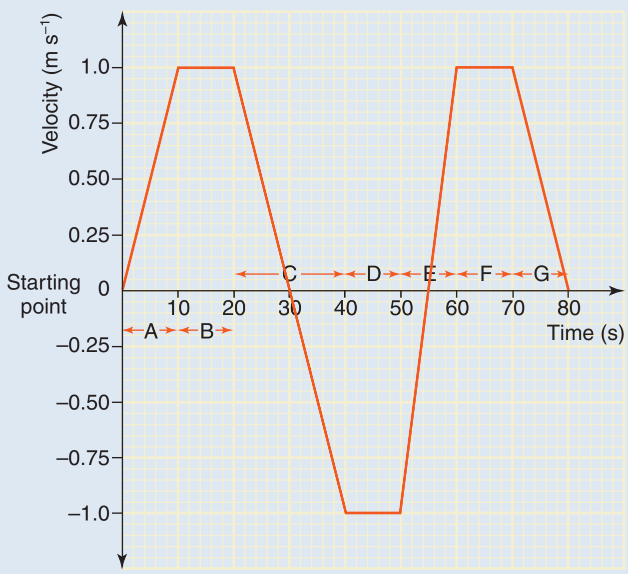


Figure 9.20

   (a) During which sections is the acceleration of the car zero?
   (b) What is the total displacement of the car during the 80 s interval?
   (c) What is the average velocity of the car during the entire interval?
   (d) At what instant did the car first reverse direction?
   (e) At what instant did the car first return to its starting point?
   (f) During which sections did the car have a negative acceleration?
   (g) During which sections was the car's speed decreasing?
   (h) Explain why your answers to (f) and (g) are different from each other.
   (i) What is the acceleration of the car throughout section E?
   (j) What is the average acceleration during the first 20 s?
   (k) Describe the motion of the remote-controlled car in words.

21. Describe in words the motion shown for each of the scenarios A, B and C in figure 9.21. Copy and complete the incomplete graphs.

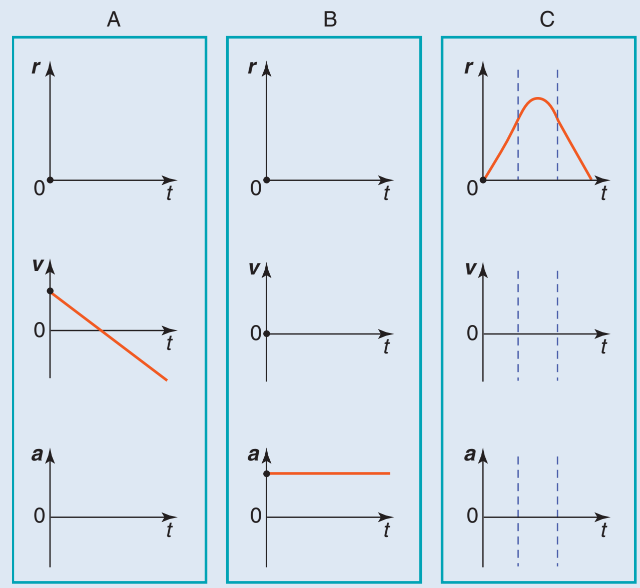


Figure 9.21

22. Figure 9.22 compares the straight-line motion of a jet ski and a car as they each accelerate from an initial speed of 5.0 m s^{-1}.

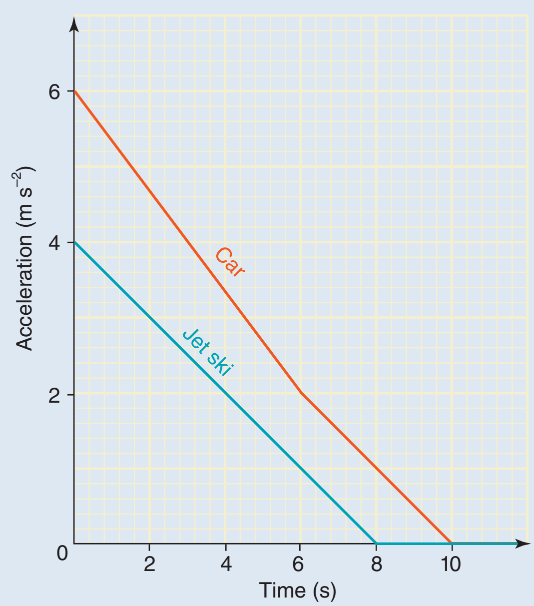


Figure 9.22

   (a) Which is the first to reach a constant speed — the jet ski or the car — and when does this occur?

(b) Calculate the final speed of:
   (i) the jet ski
   (ii) the car.

(c) Draw a graph of speed versus time describing the motion of either the jet ski or the car.

23. Once upon a time, a giant tortoise had a bet with a hare that she could beat him in a foot race over a distance of 1 km. The giant tortoise can reach a speed of about 7.5 cm s^{-1}. The hare can run as fast as 20 m s^{-1}. Both animals ran at their maximum speeds during the race. However, the hare was a rather arrogant creature and decided to have a little nap along the way. How long did the hare sleep if the result was a tie?

24. During the filming of a new movie, a stuntman has to chase a moving bus and jump into it. The stuntman is required to stand still until the bus passes him and then start chasing. The velocity-versus-time graph in figure 9.23 opposite describes the motion of the stuntman and the bus from the instant that the bus door passes the stationary stuntman.

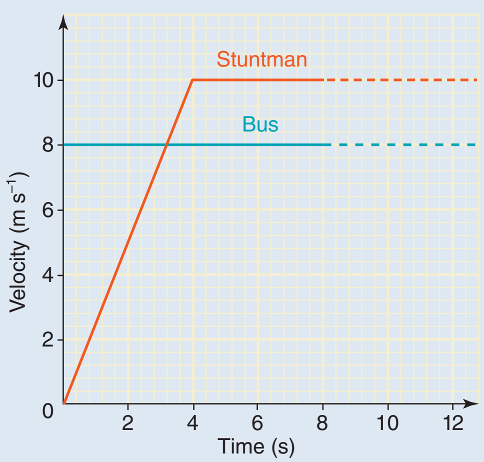


Figure 9.23

(a) At what instant did the stuntman reach the same speed as the bus?

(b) Calculate the magnitude of the acceleration of the stuntman during the first 4.0 s.

(c) At what instant did the stuntman catch up with the bus door?

(d) How far did the stuntman run before he reached the door of the bus?

25. A brand new Rolls-Royce rolls off the back of a truck as it is being delivered to its owner. It lands on its wheels. The truck is travelling along a straight road at a constant speed of 72 km h^{-1} (20 m s^{-1}). The Rolls-Royce slows down at a constant rate, coming to a stop over a distance of 240 m. It is a full minute before the truck driver realises that the precious load is missing. The driver brakes immediately, leaving a 25 m long skid mark on the road. The driver's reaction time (time interval between noticing the problem and depressing the brake) is 0.5 s. How far back is the Rolls-Royce when the truck stops?

9.1 GOING HOME

Aim

To distinguish between scalar quantities and vector quantities

Apparatus

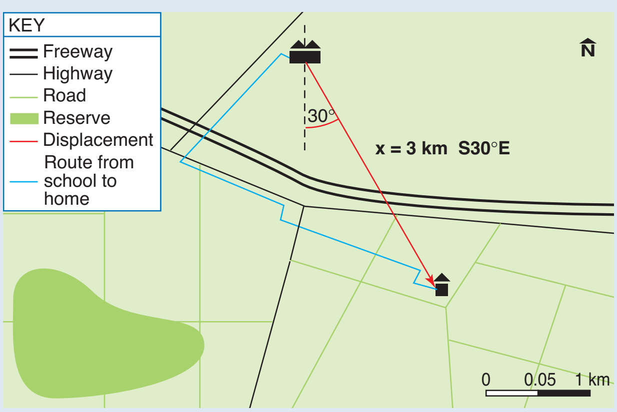
street directory or map watch

Theory

Distance and speed are scalar quantities that can be fully described as a magnitude. Displacement and velocity are vector quantities that specify magnitude and direction.

Method

Draw a map to show your journey from school to home. It should occupy about half of an A4 page and be drawn to scale. An example of a map is shown in figure 9.24. Record the time taken to travel home on a typical school day.



Time to travel home: 10 min
Displacement: 3 km S30°E
Average velocity: $\frac{3000 \text{ m}}{600 \text{ s}}$ = 5 m s^{-1} S30°E
Total distance travelled: 4.2 km
Average speed: $\frac{4200 \text{ m}}{600 \text{ s}}$ = 7 m s^{-1}

Figure 9.24 *A map showing a journey from school to home*

Results

Draw and label your displacement on the map.

Analysis and questions

Determine and specify fully:
(a) your displacement
(b) your resultant average velocity during the journey home
(c) the total distance travelled
(d) your resultant average speed during the journey home.

9.2 ON YOUR BIKE

Aim

To record the motion of a cyclist and use a graphical method to describe and analyse the motion

Apparatus

10 stop watches bicycle and helmet
100 m measuring tape speedometer

Theory

The instantaneous velocity of an object can be found from a graph of its displacement versus time by calculating the gradient of the graph. For straight-line motion in one direction only the speed is the same as the magnitude of the velocity.

The instantaneous acceleration of an object can be found from a graph of its velocity versus time by calculating the gradient of the graph.

The displacement of an object during a time interval can be found by determining the area under its velocity-versus-time graph.

Method

Record the motion of a bicycle in a straight line over a distance of 100 m. Place timekeepers at 10 m intervals along the track. The role of each timekeeper is to record the time interval between the start and the instant that the cyclist passes.

Results

Construct a table similar to the table below in which to record your results.

TIME (s)	DISPLACEMENT (m)
	0
	10
	20

Analysis and questions

1. What was the average speed (in m s^{-1}) of the cyclist?

Use the table to construct a graph of displacement versus time. Use the graph to answer the following questions.

2. What information does the gradient of the displacement-versus-time graph provide?
3. At what instant did the maximum speed occur?

4. What was the maximum speed (in m s^{-1})?
5. Express the maximum speed in km h^{-1}.

Use your displacement-versus-time graph to construct a velocity-versus-time graph of the motion. Use the velocity-versus-time graph to answer the following questions.

6. How can the acceleration be determined from your velocity-versus-time graph?
7. During which time interval was the acceleration greatest?
8. Was the acceleration zero at any time during the ride? If so, at what instant, or during which time interval, was the acceleration zero?
9. During which time interval (if any) was the acceleration negative?

Calculate the area under the velocity-versus-time graph.

10. Did you get the result that you expected? What does your result indicate about your graph?

9.3 ANALYSING MOTION WITH A CONSTANT ACCELERATION

Aim

To record the motion of a object down an inclined plane and use a graphical method to describe and analyse the motion

Apparatus

trolley or linear air-track glider
brick or other object (or objects) to raise one end of the plane
timing and recording device (e.g. ticker-timer, spark generator, photogates or motion detector and computer interface)
metre rule

Theory

If you are using ticker-timer, a spark generator or photogates to record the motion, you will need to make use of the following observation. For an object moving with a constant acceleration, the instantaneous velocity mid-way through a time interval is equal to the average velocity during that time interval.

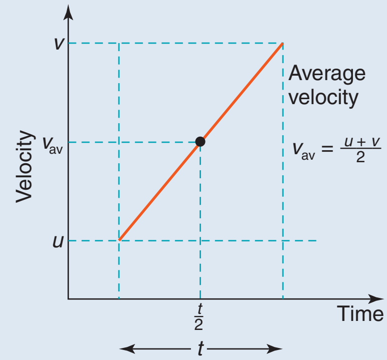


***Figure 9.25** When acceleration is constant, the instantaneous velocity at time $\frac{t}{2}$ is equal to the average velocity during the time interval t.*

Method

Make an inclined plane by raising one end of a laboratory bench or a linear air track. Use an angle of approximately 10° to the horizontal. Prepare the recording device and record the motion of a low-friction trolley or air-track glider as it accelerates down the inclined plane.

Results

If your data is recorded on ticker tape, find at least eleven consecutive clear dots.

Analysis

Use your data to determine the instantaneous velocity at enough instants of time to allow you to plot a graph of velocity versus time. Use a table to record time and instantaneous velocity. Include a third column in the table in which to record the acceleration.

Use your velocity-versus-time graph to determine the acceleration at a number of instants. Record the acceleration in your table and plot a graph of acceleration versus time.

Questions

1. What was the average acceleration of the trolley or glider?
2. Describe how the acceleration changes (if it does) while the trolley or glider moves along the inclined plane.
3. If the acceleration is not constant, explain why and suggest how the experiment could be improved so that it was constant.
4. What is the greatest source of error in measuring the instantaneous velocity of the trolley or glider?
5. How could the experiment be changed so that the error in measuring the instantaneous velocity of the trolley or glider was reduced?
6. Use your graph of velocity versus time to estimate the distance travelled by the trolley or glider. How does this distance compare with the distance measured with a metre rule?

CHAPTER 10

FORCE AND NEWTON'S LAWS OF MOTION

Figure 10.1 ***The motion of all vehicles depends on the sum of all of the forces acting on them. The same statement applies to the occupants.***

Remember

Before beginning this chapter, you should be able to:

- qualitatively describe the relationship between force, mass and acceleration
- qualitatively relate the acceleration of an object to a change in speed as a result of a net force.

Key content

At the end of this chapter you should be able to:

- use vector diagrams to solve problems involving force and acceleration
- qualitatively analyse common events involving motion in terms of Newton's laws of motion
- explain that the velocity of an object can only change if there is a non-zero net external force acting on it
- define inertia as the tendency of an object to resist a change in its motion
- explain why Newton's First Law of Motion is not apparent in many situations
- explain how the speed and/or direction of a vehicle can be changed
- describe and explain the effects of external forces — including gravity, friction, air resistance — on objects
- define mass and weight, and explain how they relate to the force of gravity
- describe the net force acting on an object
- describe the external forces that cause change in the velocity of a vehicle in a range of circumstances
- interpret Newton's Second Law of Motion and apply the equation $\Sigma \boldsymbol{F} = m\boldsymbol{a}$ to a range of situations involving modes of transport
- recall Newton's Third Law of Motion
- use the formula $F = \frac{mv^2}{r}$ to solve problems involving vehicles travelling around curves.

10.1 ANALYSING FORCES

A **force** is a push or a pull. Force is a vector quantity.

A **force** is a push or a pull. Forces can start things moving, stop them, or change their speed or direction. Some types of force require contact. For example, the force that pushes a netball or basketball towards the goal requires the contact of your hand. The friction that opposes the motion of a bicycle or car along the road requires contact of the vehicle with the road surface. However, some forces do not require contact. For example, the force of gravity pulls you down even when you are not in contact with the Earth. A magnet attracts certain materials without being in contact with them. Force is a vector quantity. In order to describe a force fully, its direction needs to be specified as well as its magnitude or size.

The SI unit of force is the newton (N). The force of gravity on a 100 g apple is about 1 N downwards. A medium-sized car starting from rest is subjected to a forward force of about 4000 N.

PRACTICAL ACTIVITIES

10.1

Force as a vector

An attraction to Earth

The apple in figure 10.2 is attracted to the Earth by the force of gravity. Even before it falls, the force of gravity is pulling it down. However, before it falls, the tree branch is also pulling it up with a force of equal magnitude.

The force of gravity is a force of attraction that exists between any pair of objects that have mass. Gravity is such a small force that, unless at least one of the objects is as massive as a planet or a natural satellite like the Moon, it is too small to measure.

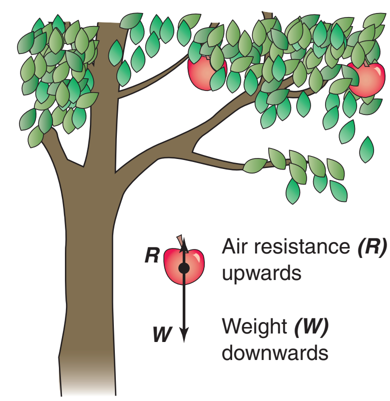


Figure 10.2 *Force is a vector quantity. Symbols representing vector quantities are in bold italic type in this text.*

Weight is the force applied to an object due to gravitational attraction.

The force on an object due to the pull of gravity is called **weight** and is usually given the symbol $\boldsymbol{W}$. The magnitude of the weight of an object is directly proportional to its mass (m). Thus,

$W \propto m$.

The weight of an object also depends on where it is. For example, the weight of your body on the Moon is considerably less than it is on Earth. Your mass remains the same wherever you are because it is a measure of the amount of matter in an object or substance. The **gravitational field strength**, which is usually given the symbol $\boldsymbol{g}$, is defined as the force of gravity on a unit of mass. Gravitational field strength is a vector quantity. In symbols,

Gravitational field strength (g) is the force of gravity on a unit of mass.

$$\boldsymbol{g} = \frac{\boldsymbol{W}}{m}.$$

Thus:

$$\boldsymbol{W} = m\boldsymbol{g}.$$

The gravitational field strength, $\boldsymbol{g}$, can be expressed in units of N kg^{-1}. However, $\boldsymbol{g}$ can also be expressed in units of m s^{-2} because it is also equal to the acceleration due to gravity. The reason for this is explained , on page 210.

The magnitude of the gravitational field strength, $\boldsymbol{g}$, at the Earth's surface is, on average, 9.81 m s^{-2}. Its magnitude decreases as altitude (height above sea-level) increases. The magnitude of $\boldsymbol{g}$ also decreases as one moves from the poles towards the equator. The magnitude of the gravitational field strength at the surface of the Moon is approximately one-sixth of that at the surface of the Earth. Table 10.1 shows the magnitude of $\boldsymbol{g}$ at several different locations on Earth.

Table 10.1 **Variation in gravitational field strength**

LOCATION	ALTITUTDE (m)	LATITUDE	MAGNITUDE OF g (m s^{-2})
Equator	0	0°	9.780
Sydney	18	34°S	9.797
Melbourne	12	37°S	9.800
Denver	1609	40°N	9.796
New York	38	41°N	9.803
North Pole	0	90°N	9.832

The magnitude of $\boldsymbol{g}$ at the Earth's surface will be taken as 9.8 m s^{-2} throughout this text. At the surface of the Moon, the magnitude of $\boldsymbol{g}$ is 1.6 m s^{-2}.

SAMPLE PROBLEM 10.1

Calculating weight

Calculate the weight of a 50 kg student:
(a) on the Earth
(b) on the Moon.

SOLUTION

(a) on the Earth

$$W = mg$$
$$= 50 \times 9.8$$
$$= 490 \text{ N downwards}$$

(b) on the moon

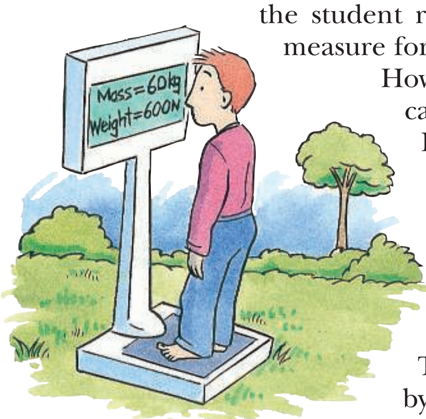
$$W = mg$$
$$= 50 \times 1.6$$
$$= 80 \text{ N downwards}$$

Note that the direction must be stated to describe the weight fully as weight is a vector.

PHYSICS FACT

Bathroom scales are designed for use only on Earth. Fortunately (at this time), that's where most of us live.

If a 60 kg student stood on bathroom scales on the Moon, the reading would be only about 10 kg. Yet the mass of the student remains at 60 kg. Bathroom scales measure force, not mass.



However, scales are designed so that you can read your mass in kilograms on Earth. Otherwise, you would have to divide the measured force by 9.8 to determine your mass. The manufacturer of the bathroom scales saves you the trouble of having to do this.

The 60 kg student has a weight of about 600 N on the Earth. However, on the Moon the student's weight is only about 100 N. The reading on the Earth-manufactured scales will be 100 N divided by 9.8 m s^{-2}, giving the result of 10 kg.

The net force

It is almost certain that at this very moment you are sitting on a chair with your feet on the floor. If your weight were the only force acting on you, what would happen? What stops you from falling through the floor?

Clearly, there must be at least one other force acting on you to stop you from falling through the floor. As figure 10.3 shows, the chair is pushing upwards on your body and the floor is pushing up on your feet. (You can actually control the size of each of these two upward forces by pushing down with your feet. However, that's another story.) The sum of these upward forces must be just enough to balance the pull of gravity downwards. Each one of these upward forces, or support forces, is called a **normal reaction** force. It is described as a *normal* force because it acts at right angles to the surface. It is described as a *reaction* force because it is only acting in response to the force that your body is applying to the floor.

The normal reaction is a force that acts perpendicular to a surface as a result of an object applying a force to the surface.

Each apple in figure 10.4 has two forces acting on it. The falling apple is speeding up as it falls because the downward force of its weight is greater than the upward force of air resistance. The vector sum of the forces acting on an object is called the **net force**.

The vector sum of the forces acting on an object is called the net force.

The direction of the net force on the apple is downwards. The apple still in the tree is not experiencing a change in its motion because the branch is pulling upwards with a force equal to its weight. The vector sum of the forces acting on it is zero.

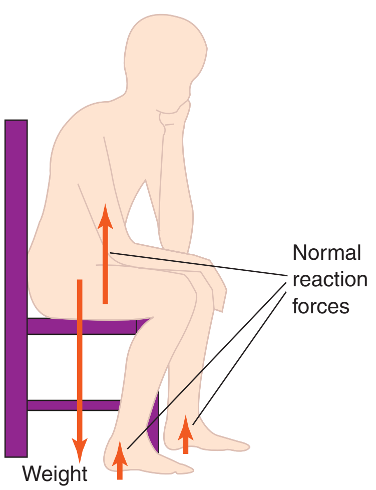


Figure 10.3 ***There is more than one force acting on you when you sit on a chair.***

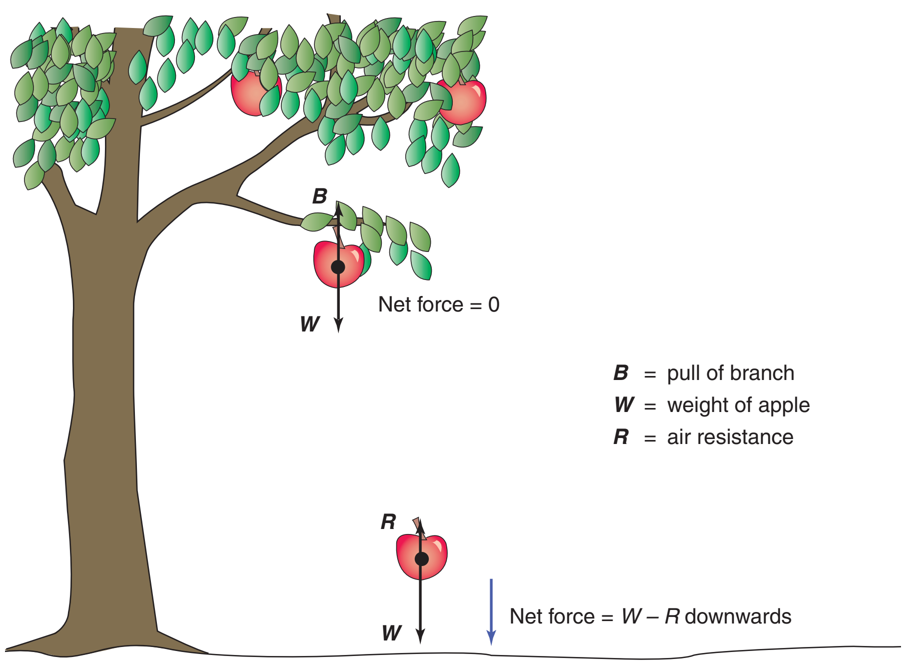


Figure 10.4 ***The motion of the apple does not change if the net force acting on it is zero.***

Adding forces together

When more than one force acts on an object, the net force is found by vector addition of the forces. If an object has two forces acting on it, one of 30 N and another of 40 N, the sum of the two forces, or net force, is 70 N only if both forces are in the same direction. The diagrams in figure 10.5 on page 200 show these two forces acting in different directions. The net force is indicated in each of the three examples. The net force is usually denoted by the symbol ΣF.

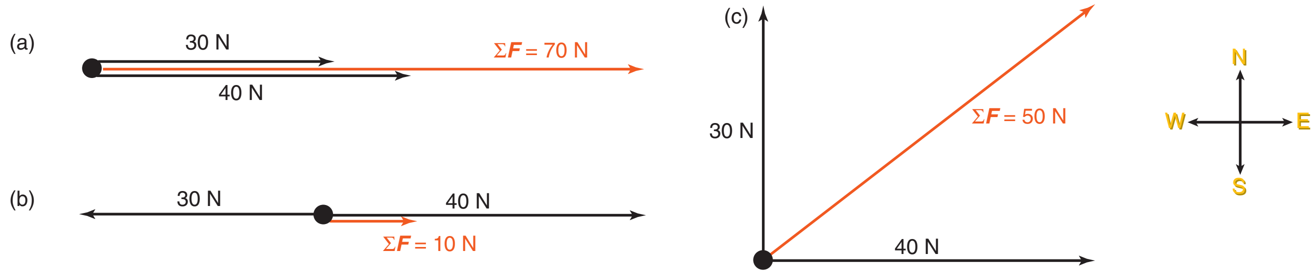


Figure 10.5 *Adding forces together*

When two or more forces are to be added together, they can be drawn to scale and then placed 'head to tail' in any order. The net force is represented by the arrow drawn from the tail of the first vector to the head of the last vector added. The order is not important as long as the magnitude and direction of each vector is not changed.

SAMPLE PROBLEM 10.2

Forces involved in a tug-of-war

In a three-way 'tug-of-war', the three teams (A, B and C) pull horizontally away from the knot joining the ropes with forces of 3000 N north, 2500 N south-west and 2800 N south-east respectively. Determine the net horizontal force exerted on the knot.

SOLUTION Figure 10.6 shows a diagram of the tug-of-war and two different ways of determining the net force on the knot. The net force is 800 N in a direction 15° east of south.

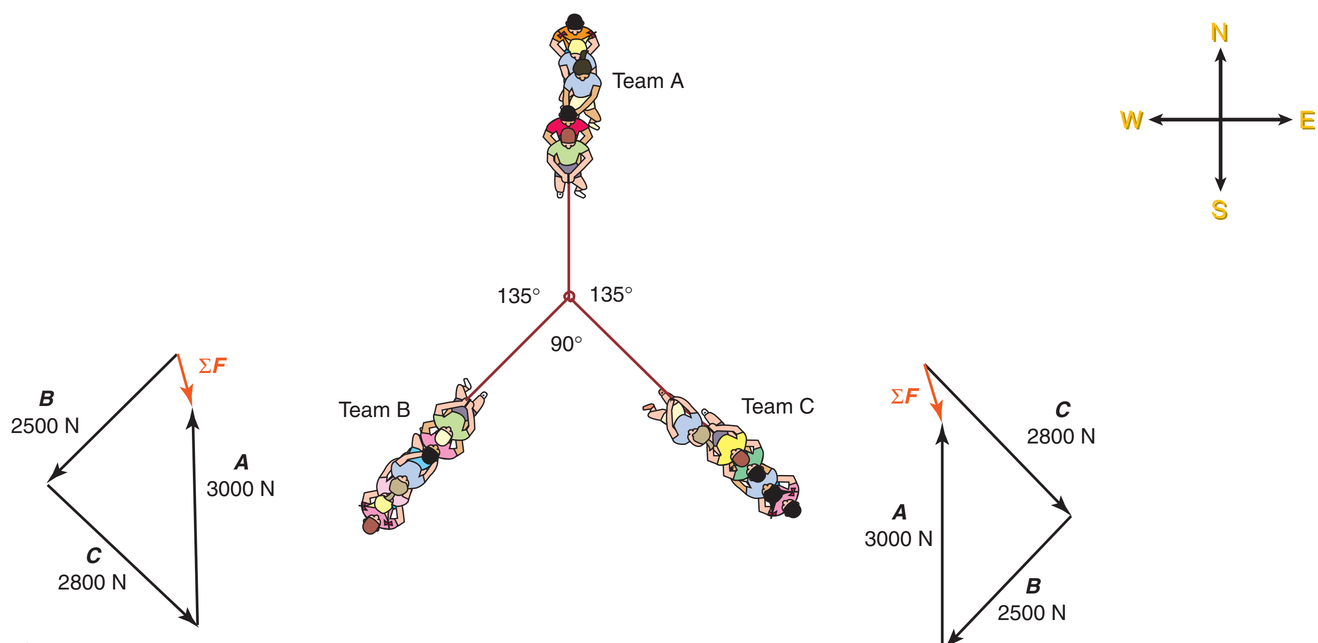


Figure 10.6

10.2 FORCES IN AND OUT OF BALANCE

When the net force on a stationary object is zero, the object remains at rest. An apple still connected to the branch of an apple tree remains at rest until the branch is no longer able to balance the weight of the apple. The net force is downwards and the motion of the apple changes. During its fall the apple continues to change its motion, accelerating until it hits the ground.

The effect of forces on the motion of objects is clearly described by Newton's First Law of Motion; that is:

Every object continues in its state of rest or uniform motion unless made to change by a non-zero net force.

Newton's First Law of Motion can also be expressed in terms of acceleration. When a non-zero net force acts on an object, it accelerates in the direction of the net force. The acceleration can take the form of a change in speed, change in direction or a change in both speed and direction.

Newton's First Law of Motion can be illustrated by flicking a coin across a tabletop. A coin flicked across a table changes its motion because the net force on it is not zero. In fact, it slows down because the direction of the net force is opposite to the direction of motion. The vertical forces, weight and the support force of the table, balance each other. The only 'unbalanced' force is that of friction. Friction is a force that surfaces exert on each other when they 'rub' together. The surface of the table applies a frictional force to the surface of the coin whenever there is an external force pushing the coin.

A coin pushed steadily across a table moves in a straight line at constant speed as long as the net force is zero (that is, as long as the magnitude of the pushing force is equal to the magnitude of the friction). The coin will speed up if you push horizontally with a force greater than the friction. It will slow down if the force of friction is greater than the horizontal pushing force. That is what happens when you stop pushing.

The ancient Greek philosopher Aristotle (384–322 BC) would have explained the motion of a coin flicked across a table by saying it stopped because there was no force present to keep it going. Aristotle believed that steady motion could only be maintained by a constant force. Aristotle's ideas about forces and motion were accepted as truth for almost 2000 years.

PHYSICS FACT

Sir Isaac Newton (1643–1727) was one of many famous scientists who were not outstanding students at school or university. He left school at 14 years of age to help his widowed mother on the family's farm. He found himself unsuited to farming, spending much of his time reading. At the age of 18, Isaac was sent to Cambridge University, where he showed no outstanding ability.

When Cambridge University was closed down in 1665 due to an outbreak of the plague, Newton went home and spent the next two years studying and writing. During this time, he developed the law of gravity, which explains the motion of the planets, and his three famous laws of motion. Newton also explained that white light consisted of many colours, and he invented calculus.

Newton's laws of gravity and motion were not published until about twenty years later. They were published in Latin in a book entitled *Philosophia Naturalis Principia*. The cost of publishing was paid by Sir Edmond Halley, the person who discovered Halley's comet.

Newton later became a member of Parliament, Warden of the Mint and President of the Royal Society. After his death in 1727, Newton was buried in Westminster Abbey, London, alongside many English kings, queens, political leaders and poets.

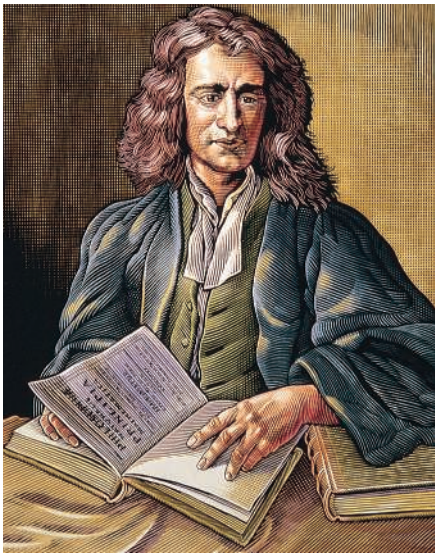


Figure 10.7 *Sir Isaac Newton*

10.3 NEWTON'S FIRST LAW AND INERTIA

Inertia is the tendency of an object to resist a change in its motion.

Newton's First Law of Motion (described in the last section) is often referred to as the Law of Inertia. The **inertia** of an object is its tendency to resist changes to its motion. Inertia is not a force; it is a property of all objects. The inertia of an object depends only on its mass. For example, a large adult on a playground swing is more difficult to get moving than a small child on the same swing. It is also more difficult to stop or change the direction of motion of a large adult than a small child.

Your inertia can be a serious problem when you are in or on a fast-moving object. As a passenger in a fast-moving car that suddenly stopped (e.g. an emergency or a collision), you would continue to move at high speed until a non-zero net force stopped you. Your inertia would resist the change in motion. The car would have inertia as well. However, it would have stopped as a result of braking or colliding with another object. If you were wearing a seat belt, it would apply a force that would stop you, along with the rest of the car. Without it, you would collide with part of the car — usually the dashboard or the windscreen. If you were to crash into a solid object while riding a bicycle, you would continue to move forward after the bicycle stopped.

Your inertia is also evident when you are in a vehicle that starts rapidly or is pushed forward in a collision. Your seat pushes your body forward along with the car. However, without a properly fitted headrest, your head remains at rest until it is pulled forward by your spine. The resulting injuries are called whiplash injuries. The purpose of a headrest is to ensure that when a large net force pushes the car forward, your head is pushed forward at the same rate as the rest of your body.

Cruising along

The forces acting on a car being driven along a straight horizontal road are shown in figure 10.8 and described below.

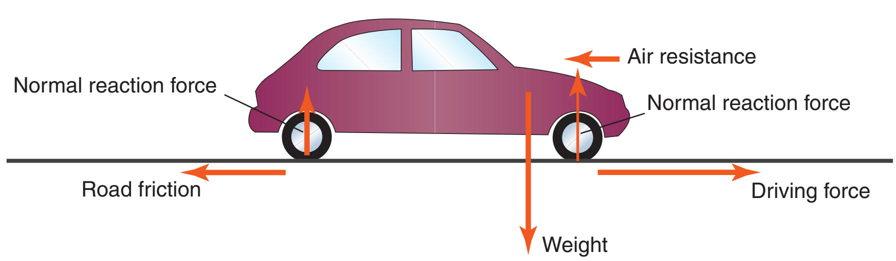


Figure 10.8 *The motion of a car on a horizontal road depends on the net force acting on it.*

The centre of mass of an object is the point at which all of its mass can be considered to be. The centre of gravity of an object is the point at which the weight can be considered to act. For an object in a uniform gravitational field, the centre of mass and the centre of gravity are at the same point.

- *Weight.* A medium-sized sedan containing a driver and passenger has a weight of about 1.5×10^4 N. The weight acts through the centre of mass, or balancing point, of the car. This is normally closer to the front of the car than the back.
- *Normal reaction force.* A normal reaction force pushes up on all four wheels. Its magnitude is usually greater at the front wheels than the rear wheels. On a horizontal road, the sum of these normal reaction forces must have the same magnitude as the weight. What do you think would happen if this was not the case?

- *Driving force*. This is provided by the road and is applied to the driving wheels. The driving wheels are turned by the motor. In most cars, either the front wheels or the rear wheels are the driving wheels. The motor of a four-wheel-drive vehicle turns all four wheels. As the tyre pushes back on the road, the road pushes forward on the tyre, propelling the car forward. The forward push of the road on the tyre is a type of friction commonly referred to as traction, or grip. If the tyres do not have enough tread, or the road is icy, there is not enough friction to push the car forward and the tyre slides on the road. The wheel spins and the car skids. The car cannot be propelled forward as effectively. Skidding also occurs if the motor turns the driving wheels too fast.
- *Road friction*. The non-driving wheels of front-wheel-drive cars roll as they are pulled along the road by the moving car. In older cars, the non-driving wheels are usually at the front. They are pushed along the road by the moving car. Rolling friction acts on the non-driving wheels in a direction opposite to the direction of movement of the car. When the driving wheels are not being turned by the motor, rolling friction opposes the forward movement of all four wheels. When the brakes are applied, the wheels to which the brakes are attached are made to turn too slowly for the speed at which the car is moving. They are no longer rolling freely. This increases the road friction greatly and the car eventually stops. If the brakes are applied hard enough the wheels stop completely, or lock, and the car goes into a skid. The sliding friction that exists when the car is skidding is less than the friction that exists when the wheels are rolling just a little.

Air resistance is a force applied to an object by the air through which it is moving.

- ***Air resistance***. The drag, or air resistance, acting on the car increases as the car moves faster. Air resistance is a form of friction that can be reduced by streamlining the vehicle. This involves shaping the vehicle so that it disturbs the air less.

The net force acting on the car in figure 10.8 is zero. It is therefore moving along the road at constant speed. We know that it is moving to the right because both the air resistance and road friction act in a direction opposite to the direction of motion. If the car were stationary, neither of these forces would be acting at all.

eBook*plus*

eLesson:
Air resistance
eles-0035

- When the driver pushes down on the accelerator, the driving force increases. The car speeds up until the sum of the air resistance and road friction grow large enough to balance it. Then, once again, the car would be moving at a constant, although higher, speed.
- When the driver stops pushing down on the accelerator, the motor stops turning the driving wheels and the driving force becomes zero. The net force would be to the left. As the car slows down, the air resistance and road friction would gradually decrease until the car comes to a stop. The net force on the car becomes zero until the driving force is restored.
- When the steering wheel is turned on a curve in the road, the direction of the driving force changes. This causes the direction of the car's velocity to change, keeps it on the road and keeps it in the correct lane. For example, when you turn the steering wheel clockwise, the wheels turn towards the right, the driving force towards the right increases and the velocity changes direction towards the right. Even though there may be no change in speed, this change in motion is an acceleration because there is a change in velocity.

PHYSICS IN FOCUS

Anti-lock brake systems

When car brakes are applied too hard, as they often are when a driver panics in an emergency, the wheels lock. The car skids, steering control is lost and the car takes longer to stop than if the wheels were still rolling. Drivers are often advised to 'pump' the brakes in wet conditions to prevent locking. This involves pushing and releasing the brake pedal in quick succession until the car stops. This, however, is very difficult to do in an emergency situation.

Anti-lock brake systems (ABSs) allow the wheels to keep rolling no matter how hard the brakes are applied. A small computer attached to the braking system monitors the rotation of the wheels. If the wheels lock and rolling stops, the pressure on the brake pads (or shoes) that stops the rotation is reduced for a very short time. This action is repeated up to 15 times each second. Anti-lock brake systems are most effective on wet roads. However, even on a dry surface, braking distances can be reduced by up to 20 per cent.

Rolling downhill

A car left parked on a hill will begin to roll down the hill with increasing speed if it is left out of gear and the handbrake is off. Figure 10.9 shows the forces acting on such a car. In order to simplify the diagram, all of the forces are drawn as if they were acting through the centre of mass of the car. The direction of net force acting on the car is down the hill. It is clear that the pull of gravity (the weight of the car), is a major contributor to the downhill motion of the car.

A component is a part. Any vector can be resolved into a number of components. When all of the components are added together, the result is the original vector.

It is often useful to divide vectors into parts called **components**. Figure 10.10 shows how the weight can be broken up, or resolved into two components — one parallel to the slope and one perpendicular to the slope. Notice that the vector sum of the components is the weight. By resolving the weight into these two components, two useful observations can be made.

1. The normal reaction force is balanced by the component of weight that is perpendicular to the surface. The net force has no component perpendicular to the road surface. This must be the case because there is no change in motion perpendicular to the slope.
2. The net force is the vector sum of the component of the weight that is parallel to the surface, and the sum of road friction and air resistance.

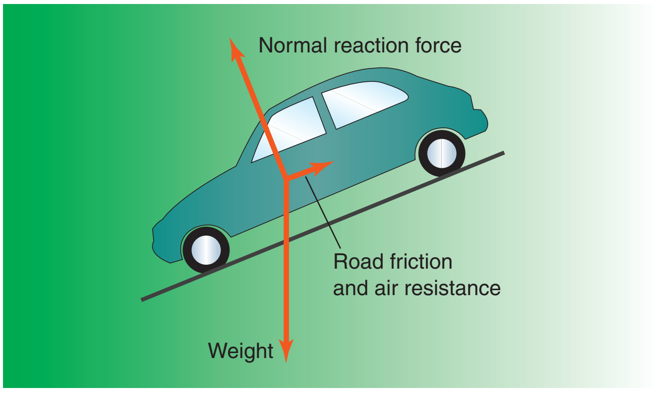


Figure 10.9 *A simplified diagram showing the forces acting on a car rolling down a slope*

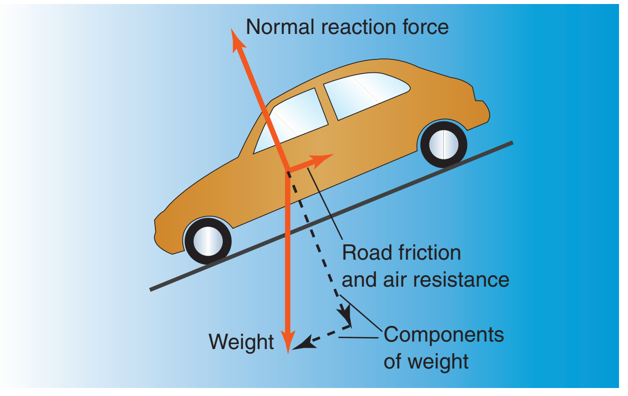


Figure 10.10 *Vectors can be resolved into components. In this case, the weight has been resolved into two components. The net force is parallel to the slope, and the car will accelerate down the slope.*

Driving uphill

When a car is driven up the slope, as shown in figure 10.11, the driving force is greater than or equal to the magnitude of the sum of the air resistance, road friction and the component of the weight that is parallel to the surface.

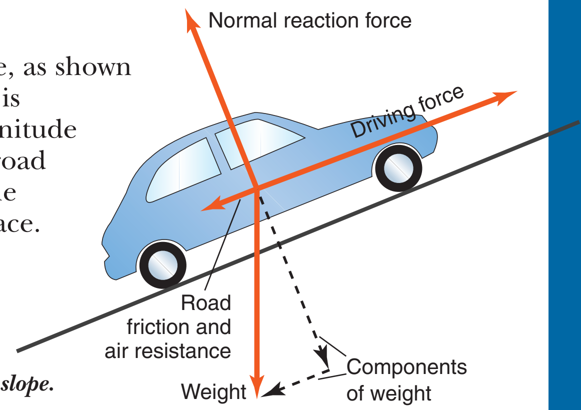


Figure 10.11 *This diagram shows the forces acting on a car driven up a slope. In this case, the car is accelerating up the slope.*

SAMPLE PROBLEM 10.3

Effect of friction on a rolling car

A car of mass 1600 kg left parked on a steep but rough road begins to roll down the hill. After a short while it reaches a constant speed. The road is inclined at 15° to the horizontal. Its speed is sufficiently slow that the air resistance is insignificant and can be ignored. Determine the magnitude of the road friction on the car while it is rolling at constant speed.

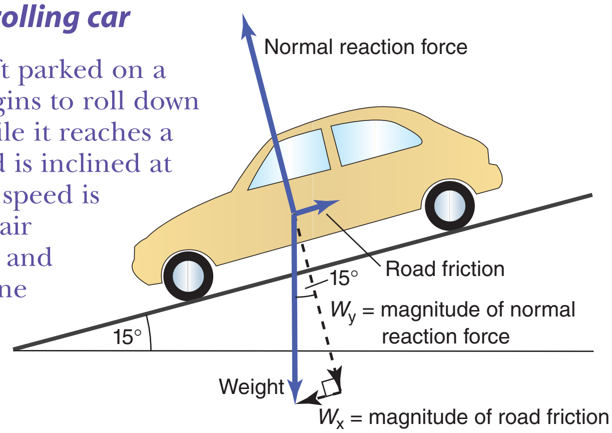


Figure 10.12

SOLUTION

Because the car is rolling at constant speed, the net force acting on it must be zero. The weight, W, can be resolved into two components — one down the slope, W_x, and one perpendicular to it, W_y. The perpendicular component of the weight, W_y, is balanced by the normal reaction force. The magnitude of the road friction must be equal to the magnitude of the weight component down the slope, W_x.

In the triangle formed by the weight and its components:

$$\sin 15° = \frac{W_x}{W}$$
$$W_x = W \sin 15°$$
$$= mg \sin 15°$$
$$= 1600 \times 9.8 \times \sin 15° \text{ (substituting data)}$$
$$= 4.1 \times 10^3 \text{ N.}$$

The magnitude of the road friction is therefore 4100 N while the car is rolling with a constant speed.

eBook*plus*

eLesson: Friction as a driving force eles-0032

Interactivity: Friction as a driving force int-0054

10.4 NEWTON'S SECOND LAW OF MOTION

Casual observations indicate that the acceleration of a given object increases as the net force on the object increases. It is also clear that lighter objects change their velocity at a greater rate than heavier objects when the same force is applied.

eBookplus

eLesson:
Newton's Second Law
eles-0033

It can be shown experimentally that the acceleration, $\boldsymbol{a}$, of an object is:

- proportional to the net force, $\Sigma \boldsymbol{F}$
- inversely proportional to the mass, m.

$$\boldsymbol{a} \propto \Sigma \boldsymbol{F} \qquad \boldsymbol{a} \propto \frac{1}{m}$$

Thus:

$$\boldsymbol{a} \propto \frac{\Sigma \boldsymbol{F}}{m}$$

$$\boldsymbol{a} = \frac{k\Sigma \boldsymbol{F}}{m}$$

where
k = a constant of proportionality.

The SI unit of force, the newton (N), is defined such that a net force of 1 N causes a mass of 1 kg to accelerate a 1 m s^{-2}. The value of the constant, k, is 1. It has no units. Thus:

$$\boldsymbol{a} = \frac{\Sigma \boldsymbol{F}}{m}$$

$$\Sigma \boldsymbol{F} = m\boldsymbol{a}. \qquad \text{(rearranging)}$$

PRACTICAL ACTIVITIES

10.2

Newton's Second Law of Motion

The equation above describes Newton's Second Law of Motion. This statement of Newton's Second Law allows you to:

- determine the net force acting on an object without knowing any of the individual forces acting on it. The net force can be deduced as long as you can measure or calculate (using formulas or graphs) the acceleration of a known mass.
- determine the mass of an object. You can do this by measuring the acceleration of an object on which a known net force is exerted.
- predict the effect of a net force on the motion of an object of known mass.

SAMPLE PROBLEM 10.4

Calculating net force

What is the magnitude of the net force acting on a 1600 kg car when its acceleration is 2.0 m s^{-2}?

SOLUTION

$$\begin{aligned} \Sigma F &= ma \\ &= 1600 \times 2.0 \\ &= 3200 \text{ N} \end{aligned}$$

SAMPLE PROBLEM 10.5

Determining the mass of a toy car

A toy car is pulled across a smooth horizontal table with a spring balance. The reading on the spring balance is 2.0 N and the acceleration of the car is measured to be 2.5 m s^{-2}. What is the mass of the toy car?

SOLUTION

Because the table is described as smooth, friction can be ignored. The force of gravity (weight) and normal reaction have no horizontal components. Because the net force is horizontal and also has no vertical component, these two vertical forces must add to zero. The net force is therefore the pull of the spring balance on the car.

$$\Sigma F = ma$$

$$2.0 \text{ N} = m \times 2.5$$

$$\begin{aligned} m &= \frac{2.0}{2.5} \\ &= 0.80 \text{ kg} \end{aligned}$$

SAMPLE PROBLEM 10.6

Using net force to determine a time interval

A 65 kg physics teacher, starting from rest, glides gracefully down a slide in the local playground. The net force on her during the slide is a constant 350 N. How long will it take her to reach the bottom of the 8.0 m slide?

SOLUTION

$$\Sigma F = ma$$
$$350\text{ N} = 65 \times a$$
$$a = \frac{350}{65}$$
$$= 5.4\text{ m s}^{-2}$$

Thus, $u = 0$, $a = 5.4\text{ m s}^{-2}$, $r = 8.0$ m and $v = ?$

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

$$8.0 = 0 + \frac{1}{2} \times 5.4 \times t^2$$
$$8.0 = 2.7\,t^2$$
$$t = \sqrt{\frac{8.0}{2.7}}$$
$$= 1.7\text{ s}$$

PHYSICS IN FOCUS

We have lift-off!

Have you ever watched a space shuttle taking off and wondered why it seems to take so long to gain speed at lift-off? Despite the huge thrust provided by its three liquid-fuel motors and two rocket boosters (resulting in an initial upward thrust of about 29 million newtons!), the initial acceleration of the space shuttle is only 3.2 m s^{-2} upwards.

Newton's Second Law of Motion provides the answer. The mass of the space shuttle at blast-off (most of which is fuel) is about 2.2 million kilograms. Its weight is

$$\mathbf{W} = m\mathbf{g}$$
$$= 2.2 \times 10^6 \times 9.8$$
$$= 2.2 \times 10^7\text{ N}$$
$$= 22\text{ million N}$$

The net force on the space shuttle at blast-off is therefore 7 million newtons up.

$$\mathbf{a} = \frac{\Sigma \mathbf{F}}{m}$$
$$= \frac{7\,000\,000}{2\,200\,000}\text{ up}$$
$$= 3.2\text{ m s}^{-2}$$

As fuel is burnt, the mass of the space shuttle decreases and the acceleration increases. Two minutes after lift-off, as the rocket boosters fall off, the acceleration is about 30 m s^{-2}.

Figure 10.13 *Why does the space shuttle seem to take so long to gain speed at lift-off?*

SAMPLE PROBLEM 10.7

Forces on a water-skier

A water-skier of mass 80 kg, starting from rest, is pulled along in a northerly direction by a horizontal rope with a constant tension of 240 N. After 6.0 s, he has reached a speed of 12 m s^{-1}.

(a) What is the net force on the skier?

(b) If the tension in the rope were the only horizontal force acting on the skier, what would his acceleration be?

(c) What is the sum of the resistance forces on the skier?

SOLUTION A diagram (figure 10.14) is used to show the forces acting on the skier. Assign the positive direction as north.

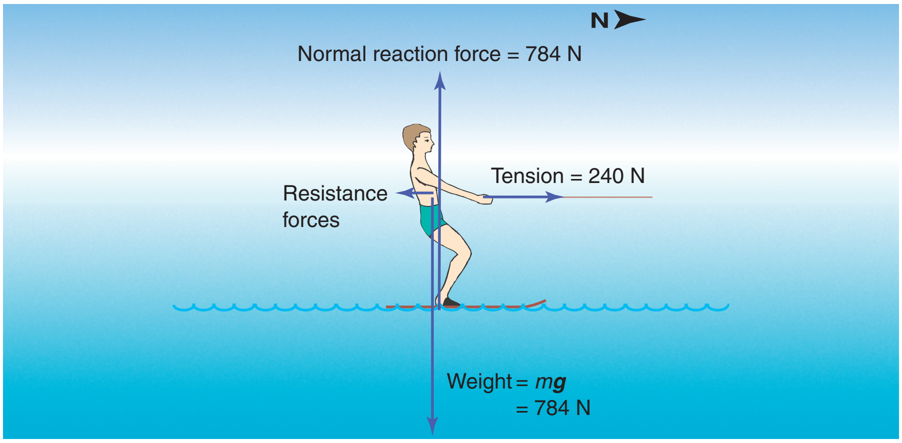


Figure 10.14

(a) The net force cannot be determined by adding the individual force vectors because the resistance forces are not given. There is no information in the question to suggest that they can be ignored. However, if the acceleration is known, the net force can be calculated by applying Newton's Second Law.

$u = 0,\ v = 12 \text{ m s}^{-1},\ t = 6.0 \text{ s},\ a = ?$

$$v = u + at$$
$$12 = 0 + a \times 6.0$$
$$a \times 6.0 = 12$$
$$a = 2.0 \text{ m s}^{-2}$$

Thus:

$$\Sigma \boldsymbol{F} = m\boldsymbol{a}$$
$$= 80 \times 2.0$$
$$= 160 \text{ N north.}$$

(b) The net force on the skier is horizontal. If the tension were the only horizontal force acting on the skier, it would be equal to the net force since the vertical forces on the skier add to zero.

Thus, the acceleration would be given by:

$$\boldsymbol{a} = \frac{\Sigma \boldsymbol{F}}{m}$$
$$= \frac{240}{80}$$
$$= 3.0 \text{ m s}^{-2} \text{ north.}$$

(c) The sum of the resistance forces (friction caused by the water surface and air resistance) on the skier is the difference between the net force and the tension.

Sum of resistance forces $= \Sigma \boldsymbol{F} - \text{tension}$
$= 160 \text{ N north} - 240 \text{ N north}$
$= 80 \text{ N south}$

SAMPLE PROBLEM 10.8

Speed and distance calculations

A loaded supermarket shopping trolley with a total mass of 60 kg is left standing on a footpath which is inclined at an angle of 30° to the horizontal. As the tired shopper searches for his car keys, he fails to notice that the loaded shopping trolley is beginning to roll away. It rolls in a straight line down the footpath for 9.0 s before it is stopped by an alert (and very strong) supermarket employee. Find:

(a) the speed of the shopping trolley at the end of its roll

(b) the distance covered by the trolley during its roll.

Assume that the footpath exerts a constant friction force of 270 N on the runaway trolley.

SOLUTION

A diagram must be drawn to show the three forces acting on the shopping trolley. Air resistance is not included as it is negligible. The forces should be shown as acting through the centre of mass of the loaded trolley as in figure 10.15. The components of the weight, which are parallel and perpendicular to the footpath surface, should also be shown on the diagram.

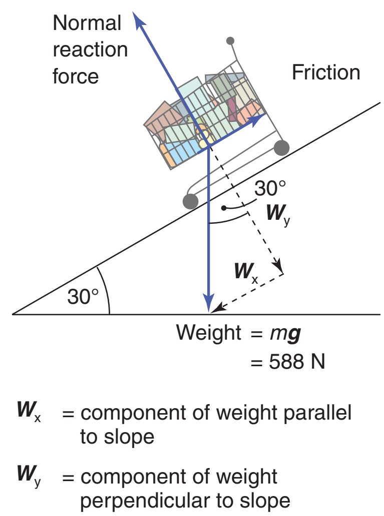


Figure 10.15

eBook*plus*

eLesson: Motion down an inclined plane eles-0034

The motion of the runaway shopping trolley, originally at rest, can be described by using the information provided, along with Newton's Second Law, which is used to determine its acceleration.

The net force can be found by 'breaking up' the weight into two components — one parallel to the footpath surface ($\boldsymbol{W}_x$) and the other perpendicular to the surface ($\boldsymbol{W}_y$). We know that $\boldsymbol{W}_y$ is balanced by the normal reaction force because there is clearly no acceleration of the trolley perpendicular to the surface. The net force is therefore down the slope and has a magnitude of:

$$\begin{aligned}\Sigma F &= W_x - \text{friction}\\ &= mg \sin 30° - 270\\ &= 588 \sin 30° - 270\\ &= 24 \text{ N.}\end{aligned}$$

Newton's Second Law can now be applied to determine the acceleration of the trolley down the slope. Assign the positive direction as down the slope.

$$\begin{aligned}\Sigma \boldsymbol{F} &= m\boldsymbol{a}\\ 24 &= 60 \times \boldsymbol{a}\\ \boldsymbol{a} &= \frac{24}{60}\\ &= 0.40 \text{ m s}^{-2} \text{ down slope}\end{aligned}$$

The final speed and distance travelled by the trolley can now be calculated.

$u = 0,\ a = 0.40 \text{ m s}^{-2},\ t = 9.0 \text{ s},\ v = ?,\ r = ?$

$$\begin{aligned}v &= u + at\\ &= 0 + 0.40 \times 9.0\\ &= 3.6 \text{ m s}^{-1}\end{aligned}$$

$$r = ut + \frac{1}{2}at^2$$
$$= 0 \times 9.0 + \frac{1}{2} \times 0.40 \times (9.0)^2$$
$$= \frac{1}{2} \times 0.40 \times 81$$
$$= 16 \text{ m}$$

At the end of its roll, the trolley was travelling at a speed of 3.6 m s^{-1} and had moved a distance of 16 m down the slope.

Falling down

Objects that are falling (or rising) through the air are generally subjected to two forces — weight and air resistance. The weight of the object is constant. The magnitude of the air resistance, however, is not constant. It depends on many factors, including the object's speed, surface area and density. It also depends on the density of the body of air through which the object is falling. The air resistance is always opposite to the direction of motion. The net force on a falling object of mass m and weight $\boldsymbol{W}$ can therefore be expressed as:

$\Sigma \boldsymbol{F} = m\boldsymbol{a}$ (where $\boldsymbol{a}$ is the acceleration of the object)

$\boldsymbol{W} -$ air resistance $= m\boldsymbol{a}$.

When dense objects fall through small distances near the surface of the Earth it is usually quite reasonable to assume that the air resistance is negligible. Thus:

$\boldsymbol{W} = m\boldsymbol{a}$

$m\boldsymbol{g} = m\boldsymbol{a}$ (where $\boldsymbol{g}$ is the gravitational field strength)

$\boldsymbol{g} = \boldsymbol{a}$.

The units N kg^{-1} and m s^{-2} are equivalent.

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

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

(multiplying both sides by kg^{-1})

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

The acceleration of a body in free fall in a vacuum or where air resistance is negligible is equal to the gravitational field strength. At the Earth's surface, where $\boldsymbol{g} = 9.8$ N kg^{-1}, this acceleration is 9.8 m s^{-2}.

If a bowling ball, a tennis ball and a table-tennis ball were dropped at the same instant from a height of 2.0 m in a vacuum, they would all reach the ground at the same time. This is because each ball would have the same initial velocity of zero and the same acceleration.

If, however, the balls are dropped either in a classroom or outside, the table-tennis ball will reach the ground a moment later than the other two balls.

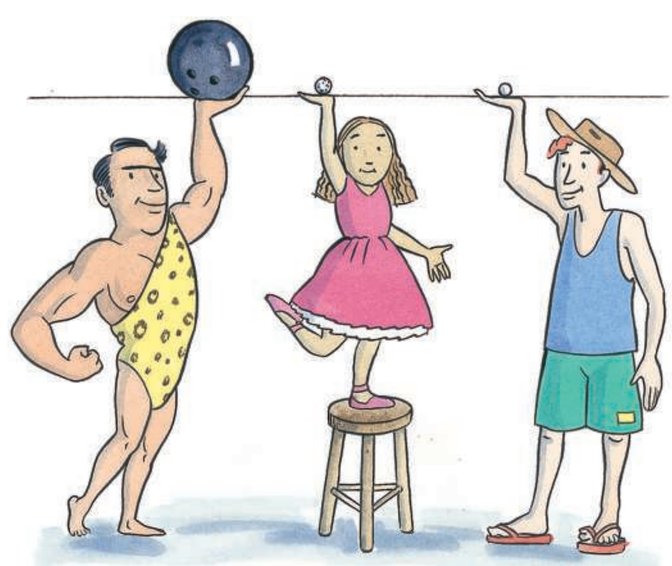


Figure 10.16 *A bowling ball, a golf ball and a table-tennis ball dropped from a height of 2.0 metres. Which one would you expect to reach the ground first?*

The acceleration of each of the balls is:

$$\begin{aligned} \boldsymbol{a} &= \frac{\Sigma \boldsymbol{F}}{m} \\ &= \frac{m\boldsymbol{g} - \boldsymbol{A}}{m} \qquad \text{(where } \boldsymbol{A} \text{ is air resistance)} \\ &= \frac{m\boldsymbol{g}}{m} - \frac{\boldsymbol{A}}{m} \\ &= \boldsymbol{g} - \frac{\boldsymbol{A}}{m}. \end{aligned}$$

The acceleration depends on the air resistance and the mass of each ball as well as $\boldsymbol{g}$.

The term $\frac{\boldsymbol{A}}{m}$ is very small for the bowling ball and the golf ball. Even though the air resistance on the table-tennis ball is small, its mass is also small and the term $\frac{\boldsymbol{A}}{m}$ is not as small as it is for the other two balls.

WARNING: Do not drop a bowling ball. If you wish to try this experiment, replace the bowling ball with a medicine ball and keep your feet out of the way!

10.5 NEWTON'S THIRD LAW OF MOTION

Forces always act in pairs (see figure 10.17). When you lift a heavy school bag you can feel it pulling down on you. When you slump into a comfortable chair at the end of a long day at school you can feel it pushing up on you. When you catch a fast-moving ball you can feel it push on your hand as you apply the force to stop it.

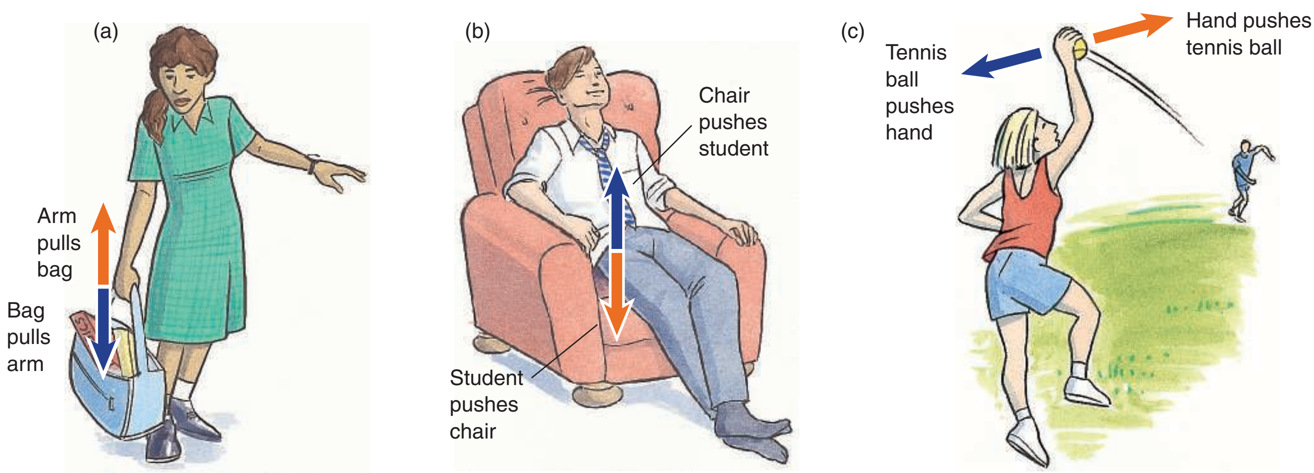


Figure 10.17 *Forces always act in pairs. (a) The arm pulls up on the bag; the bag pulls down on the arm. (b) The student pushes down and back on the chair; the chair pushes up and forward on the student. (c) The hand pushes on the ball; the ball pushes on the hand.*

Sir Isaac Newton recognised that forces always acted in pairs in his Third Law of Motion, which is most commonly stated as:

For every action there is an equal and opposite reaction.

eBook*plus*

eLesson:
Newton's laws
eles-0036
Interactivity:
Newton's laws
int-0055

A more precise statement of the Newton's Third Law is:

Whenever an object applies a force (an action) to a second object, the second object applies an equal and opposite force (called a reaction) to the first object.

It is very important to remember that the forces that make up action–reaction pairs act on different objects. That is why it makes no sense to add them together so that they 'cancel out'. The motion of each object in figure 10.17 is determined by the net force acting on it.

The net force on the student sitting in the chair in figure 10.17b is zero because the upward push of the chair is balanced by the downward force of gravity, or weight of the student.

The tennis ball in figure 10.17c slows down because the net force on the tennis ball is not zero. The push of the hand on the ball is much larger than any of the other forces acting on the ball. The net force on the hand is zero if the hand does not change its motion during the catch. The push of the ball is balanced by the push of arm muscles on the hand.

Newton's Third Law of Motion in action

The rowing boat shown in figure 10.18 is propelled forward by the push of water on the oars. As the face of each oar pushes back on the water, the water pushes back with an equal and opposite force on each oar. The push on the oars, which are held tightly by the rowers, propels them and their boat forward. A greater push (or action) on the water results in a greater push (or reaction) on the oar.

In fact, none of your forward motion, whether you are on land, water or in the air, could occur without an action–reaction pair of forces.

- When you swim, you push the water backwards with your hands, arms and legs. The water pushes in the opposite direction, propelling you forwards.
- In order to walk or run, you push your feet backwards and down on the ground. The ground pushes in the opposite direction, pushing forwards and up on your feet.
- The forward driving force on the weels of a car is the result of a push back on the road by the wheels.
- A jet or a propeller-driven plane is thrust forwards by air. The jet engines or propellers are designed to push air backwards with a very large force. The air pushes forward on the plane with an equally large force.

Figure 10.18 ***This rowing team relies on a reaction force to propel itself forward.***

Multiple bodies

Newton's Second Law of Motion can be applied to a whole system of objects or each individual object in a system.

Figure 10.19 shows a small dinghy being pulled by a larger boat. The forces acting on the larger boat are labelled in red while the forces acting on the small dinghy are labelled in green. Newton's Second Law of Motion can be applied to each of the two boats. Figure 10.20 shows only the forces acting on the system of the two boats and the rope joining them. When Newton's Second Law of Motion is applied to the whole system, the system is considered to be a single object.

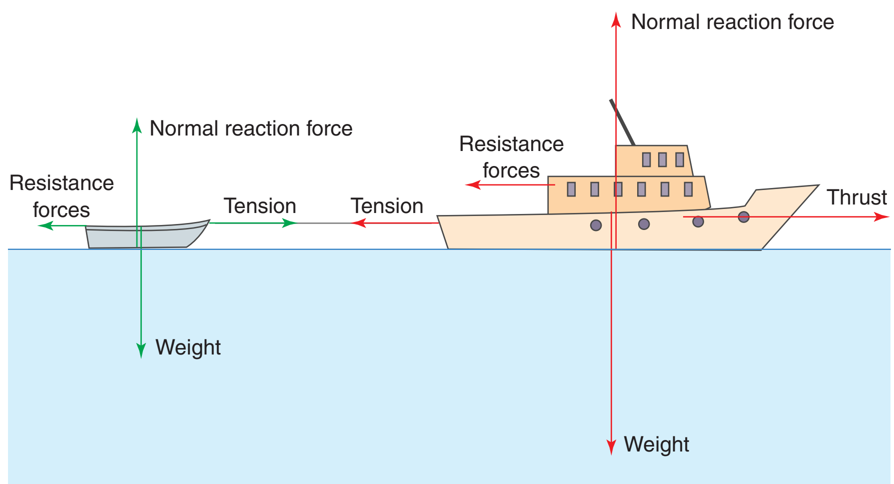


Figure 10.19 *This diagram shows the forces acting on each of the two boats.*

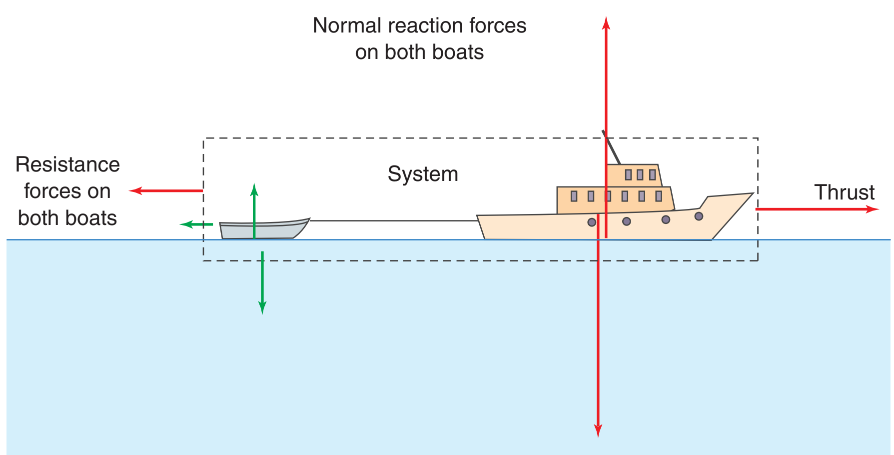


Figure 10.20 *This diagram shows the forces acting on the system. The system consists of the two boats and the rope joining them.*

Newton's Third Law of Motion is discussed again in chapter 11 in relation to the conservation of momentum.

The thrust which acts on the larger boat and the system is provided by the water. The propeller of the larger boat pushes back on the water and the water pushes forward on the propeller blades. The only force that can cause the small dinghy to accelerate forward is the tension in the rope. If the tension in the rope is greater than the resistance forces on the dinghy, it will accelerate. If the tension in the rope is equal to the resistance forces on the dinghy, it will move with a constant velocity. If the tension in the rope is less than the resistance forces on the dinghy, it will slow down. That is, its acceleration will be negative.

The rope pulls back on the larger boat with the same tension that it applied in a forward direction on the small dinghy. This is consistent with Newton's Third Law of Motion. Through the rope, the larger boat pulls forward on the small dinghy with a force that is equal and opposite to the force with which the smaller dinghy pulls on the larger boat.

SAMPLE PROBLEM 10.9

Forces on a car and trailer

A car of mass 1600 kg tows a trailer of mass 400 kg. The coupling between the car and trailer is rigid. The driving force acting on the car as it starts from rest is 5400 N in an easterly direction. The frictional forces resisting the motion of the car and trailer are insignificant and can be ignored. Calculate:

(a) the acceleration of the car and trailer
(b) the net force acting on the trailer
(c) the force applied on the trailer by the car
(d) the force applied on the car by the trailer.

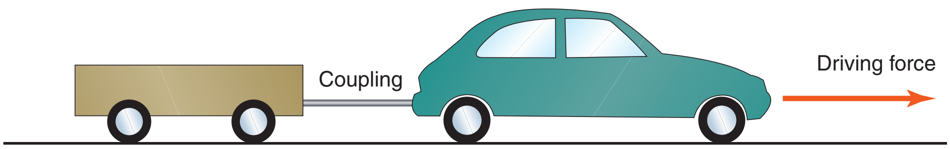


Figure 10.21 *The car towing a trailer. The only external horizontal force is the driving force.*

SOLUTION

(a) Because the coupling between the car and trailer is rigid, they have equal accelerations. Newton's Second Law of Motion can be applied to the system of the car and trailer.

$$\Sigma \boldsymbol{F} = m\boldsymbol{a}$$
$$\boldsymbol{a} = \frac{\Sigma \boldsymbol{F}}{m}$$
$$= \frac{5400}{2000}$$
$$= 2.7 \text{ m s}^{-2} \text{ east}$$

(b) Apply Newton's Second Law of Motion to the trailer.

$$\Sigma \boldsymbol{F} = m\boldsymbol{a}$$
$$= 400 \times 2.7$$
$$= 1080 \text{ N east}$$

(c) The only horizontal force acting on the trailer is the force applied by the car. The force applied on the trailer by the car is therefore 1080 N east.

(d) According to Newton's Third Law of Motion, the force applied on the car by the trailer is equal and opposite in direction to the force applied on the trailer by the car. That force is therefore 1080 N west.

Circular motion

To move in a circle, or even part of a circle, at constant speed, requires a constant net force towards the centre of the circle. This can be best illustrated by swinging a rubber stopper tied to a string in a circle above your head. Of course, this should be done outdoors, with no other person within range of the rubber stopper.

Figure 10.22 *Motion in a circle at constant speed requires a constant net force towards the centre of the circle. In this case, it is provided by keeping the tension in the string constant.*

Even though the object is moving at constant speed, it is accelerating in the same direction as the net force — towards the centre of the circle. The instantaneous acceleration of an object can be calculated over a very small time interval using the formula:

$$\boldsymbol{a} = \frac{\Delta \boldsymbol{v}}{\Delta t}$$

Centripetal acceleration is the acceleration of an object travelling in a circular path with constant speed. It is directed towards the centre of the circle.

This acceleration is called the **centripetal acceleration**. The word 'centripetal' comes from Latin words meaning 'centre-seeking'. The first task is to determine $\Delta \boldsymbol{v}$ using the vector diagram shown in figure 10.23. It shows an object changing direction by a small angle $\Delta\theta$ during the time interval Δt as it moves from A to B.

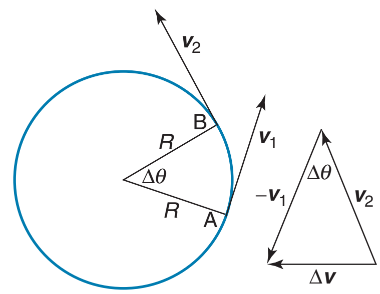


Figure 10.23 *Motion at a constant speed, v, in a circle of radius R*

eBook*plus*

Weblink:
Circular motion applet

Figure 10.24 is another vector diagram showing the displacement of the object during the same small time interval. The distance travelled along the arc AB is equal to $v\Delta t$, since speed $= \dfrac{\text{distance}}{\text{time}}$. For a small angle $\Delta\theta$, the length of the arc AB can be considered to be equal to the length of the chord AB, which is Δr. Therefore $\Delta r = v\Delta t$.

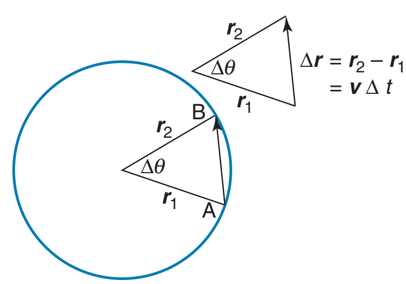


Figure 10.24 *The displacement is the change in position of the object during the small time interval.*

The triangles shown in figures 10.23 and 10.24 are similar triangles. Therefore the ratios of the magnitudes of their corresponding sides are equal. The magnitude of each of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is v. The magnitude of each of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ is r.

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

$$\frac{\Delta v}{\Delta t} = \frac{v^2}{r}$$

Therefore the magnitude of the acceleration of an object travelling at a constant speed v in circular motion of radius r can be expressed as:

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

Centripetal force is the net force on an object travelling in a circular path at constant speed. It is directed towards the centre of the circle.

The magnitude of the net force on the object, known as the **centripetal force**, can therefore be expressed as:

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

where m is the mass of the object.

SAMPLE PROBLEM 10.10

Calculating acceleration and net force around a curve

A car of mass 1200 kg is driven at a constant speed of 15 m s^{-1} around a curve with a radius of 12 m. Calculate:
(a) the magnitude of acceleration of the car
(b) the magnitude of net force acting on the car.

SOLUTION

(a) $m = 1200$ kg, $v = 15$ ms^{-1}, r = 12 m

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

$$a = \frac{15^2}{12}$$

$$a = 18.75 \text{ m s}^{-2}$$

The acceleration of the car is 19 m s^{-2}.

(b) $$\Sigma F = ma = \frac{mv^2}{r}$$

$$= 1200 \times 18.75$$

$$= 2.25 \times 10^4 \text{ N}$$

The magnitude of the net force acting on the car is 2.3×10^4 N.

Going around the bend

When a vehicle travels around a bend, or curve, at constant speed, its motion can be considered to be part of a circular motion. The curve makes up the arc of a circle. In order for a car to travel around a corner safely, the net force acting on it must be towards the centre of the circle.

Figure 10.25a shows the forces acting on a vehicle of mass m travelling around a curve with a radius, r, at a constant speed, v. The forces acting on the car are weight, $\boldsymbol{W}$, friction and the normal reaction, $\boldsymbol{N}$. The only force with a component towards the centre of the circle is the 'sideways' friction. This sideways friction makes up the whole of the magnitude of the net force (and therefore the centripetal force) on the vehicle.

$$\Sigma F = \text{sideways friction} = \frac{mv^2}{r}$$

If you drive the vehicle around the curve with a speed so that $\frac{mv^2}{r}$ is greater than the sideways friction, the motion is no longer circular and the vehicle will skid off the road. If the road is wet, sideways friction is less and a lower speed is necessary to drive safely around the curve.

If the road is banked at an angle θ towards the centre of the circle, a component of the normal reaction $N \sin \theta$ can also contribute to the centripetal force. This is shown in figure 10.25b.

Banking the road increases the net force towards the centre of the circle, and therefore the centripetal force. This means that, for a given curve, banking the road makes a higher speed possible.

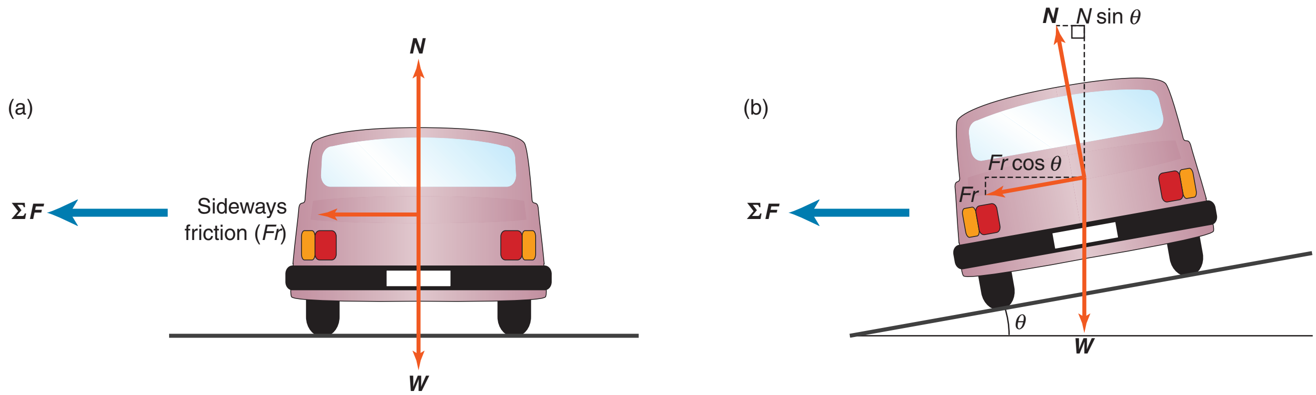


Figure 10.25 *(a) For the vehicle to take the corner safely, the net force must be towards the centre of the circle. (b) Banking the road allows a component of the normal reaction to contribute to the centripetal force.*

SAMPLE PROBLEM 10.11

Calculating maximum constant speed on a curve

A car of mass 1280 kg travels around a bend with a radius of 12.0 metres. The total sideways friction on the wheels is 16 400 N. The road is not banked. Calculate the maximum constant speed at which the car can be driven around the bend without skidding off the road.

SOLUTION

The car will maintain the circular motion around the bend if:

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

where v = maximum speed

If v were to exceed this speed $\Sigma F < \frac{mv^2}{r}$, the circular motion could not be maintained and the vehicle would skid.

$$\Sigma F = \text{sideways friction} = 16\,400 = 1280 \times \frac{v^2}{12.0}$$

$$v^2 = 16\,400 \times \frac{12.0}{1280} = 153.75$$

$$v = 12.4 \text{ s}^{-1}$$

The maximum constant speed at which the vehicle can be driven around the bend is 12.4 m s^{-1}.

SUMMARY

- Force is a vector quantity.
- Weight is a measure of the force on an object due to the pull of gravity.
- The weight of an object is directly proportional to its mass.
- The vector sum of the forces acting on an object is called the net force.
- The velocity of an object can only change if there is a non-zero net force acting on it. This statement is an expression of Newton's First Law of Motion.
- When a non-zero net force acts on an object, it accelerates in the direction of the net force.
- Acceleration occurs when there is a change in speed and/or direction.
- Inertia is the tendency of an object to resist a change in its motion.
- The forces acting on a moving vehicle are:
  - weight, downwards
  - the normal reaction force, applied perpendicular to the surface of the road
  - the driving force, applied in the direction of motion by the road
  - road friction, applied to the non-driving wheels opposite to the direction of motion
  - air resistance, applied opposite to the direction of motion.
- The motion of a vehicle depends on the net force acting on the vehicle.
- Newton's Second Law of Motion describes the relationship between the acceleration of an object, the net force acting on it, and the object's mass. It can be expressed as $\Sigma \boldsymbol{F} = m\boldsymbol{a}$.
- Newton's Second Law can be applied to a single object, or a system of multiple bodies which are in contact or connected together.
- When an object applies a force (an action) to a second object, the second object applies an equal and opposite force (a reaction) to the first object. This statement is an expression of Newton's Third Law of Motion.
- An object moves in a circle with constant speed if it experiences a net force of constant magnitude towards the centre of the circle. This net force towards the centre of the circle is called the centripetal force.
- The magnitude of the centripetal force acting on an object moving in a circle, or part of a circle, with constant speed is given by the equation $\Sigma F = \frac{mv^2}{r}$.

QUESTIONS

Assume that the magnitude of the gravitational field strength at the Earth's surface is 9.8 m s^{-2}.

1. Describe the difference between a vector quantity and a scalar quantity.
2. State which of the following are vector quantities:
   (a) mass
   (b) weight
   (c) gravitational field strength
   (d) time
   (e) acceleration.
3. A slightly overweight physics teacher steps off the bathroom scales and proudly remarks 'My weight is down to 75 kg!' The physics teacher clearly should have known better. Rewrite the remark in two different ways so that it is correct.
4. A Mitsubishi Magna sedan has a mass of 1400 kg with a full tank of petrol.
   (a) Calculate the magnitude of its weight at the surface of the Earth.
   (b) Calculate the weight of the car on the surface of Mars where the magnitude of the gravitational field strength is 3.6 m s^{-2}.
   (c) Calculate the mass of the Mitsubishi Magna on the surface of Mars.
5. Estimate your own mass in kilograms and calculate:
   (a) the magnitude of your weight at the surface of the Earth
   (b) your weight on the surface of Mars where the magnitude of the gravitational field strength is 3.6 m s^{-2}
   (c) your mass on the planet Mars.
6. The set of kitchen scales in figure 10.26a is used to determine mass. As the spring inside is compressed, the pointer in front of the scale moves. The beam balance in figure 10.26b is used in many school laboratories to determine mass. Which of the two instruments would you prefer to use to measure the mass of a small rock (with a mass of less than 300 grams) on the Moon? Explain your answer.

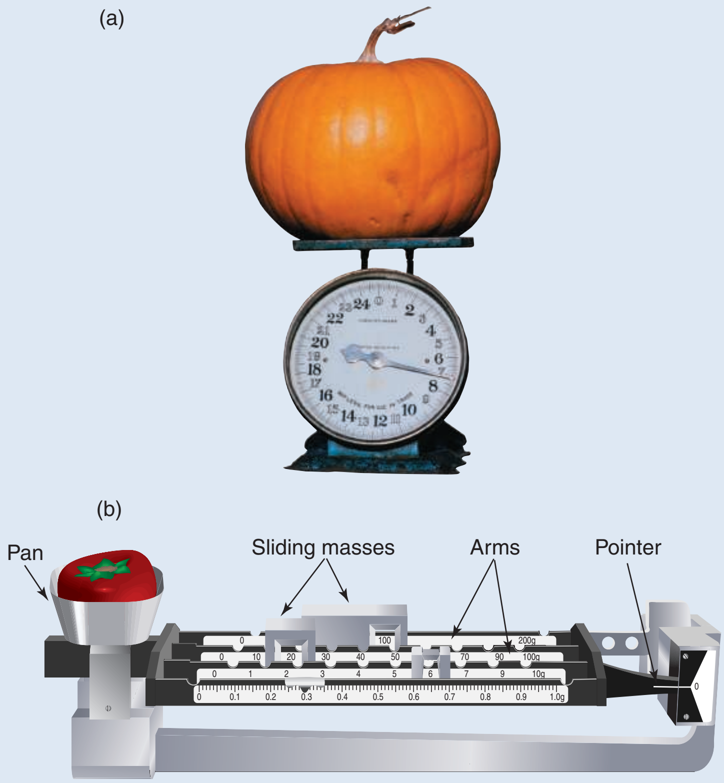


Figure 10.26

7. Determine the net force in each of the situations illustrated in figure 10.27.

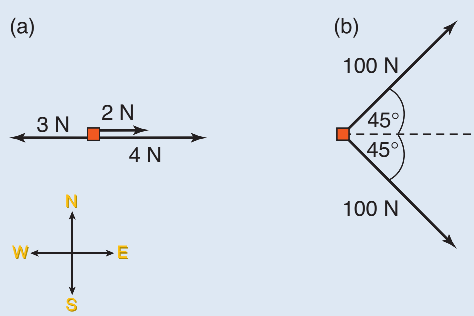


Figure 10.27

8. In the illustrations in figure 10.28, the net force is shown along with all but one of the contributing forces. Determine the magnitude and direction of the missing force.

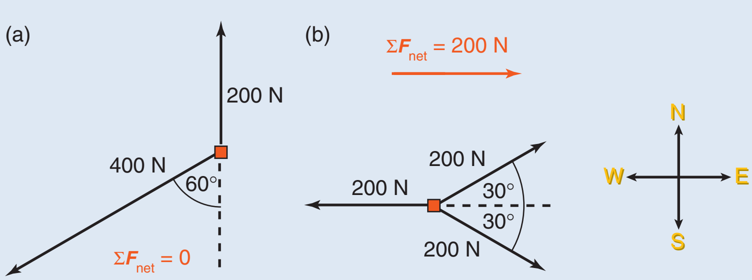


Figure 10.28

9. When you stand in an elevator there are only two significant forces acting on you — your weight and the normal reaction force. It is important to note that the tension in the cable is not pulling on you — it is pulling on the elevator. The only object that can push you upwards is the floor of the elevator.

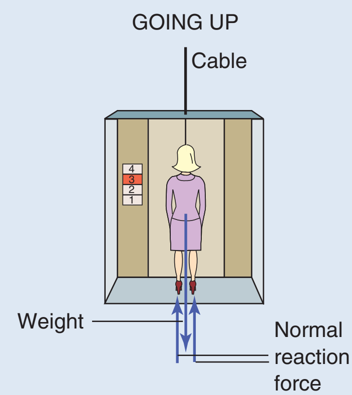


Figure 10.29 *The forces acting on you in an elevator*

   (a) State whether the normal reaction force is less than, equal to or greater than your weight when the elevator is:
      (i) stationary
      (ii) moving upwards with a constant speed
      (iii) speeding up on its way to the top floor
      (iv) slowing down as it approaches the top floor.
   (b) Explain how the movement of elevators in tall buildings sometimes makes you feel 'heavy' or 'light'.
10. A car is moving north on a horizontal road at a constant speed of 60 km h^{-1}.
   (a) Draw a diagram showing all of the significant forces acting on the car. Show all of the forces as if they were acting through the centre of mass.
   (b) Calculate the net force on the car.
11. When you are standing on a bus or train that stops suddenly, you lurch forwards. Apply Newton's First Law of Motion to explain why this happens.
12. The ancient Greek philosopher Aristotle would have explained a car rolling to a stop on a horizontal road by saying that it slowed down because there was no constant force to keep it going. Propose a better explanation.
13. If the bicycle that you are riding runs into an obstacle like a large rock, you may be flung forwards over the handlebars. Explain in terms of inertia why this happens.

14. When you try to push a broken-down car with its handbrake still on, it does not move. Explain other forces that are acting on the car to produce a net force of zero.
15. Explain why a car takes longer to stop if the brakes are applied too hard.
16. Determine the magnitude of the horizontal components of each of the following forces (figure 10.30).

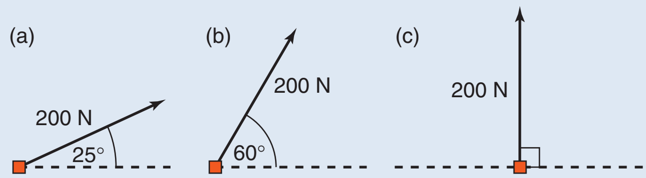


Figure 10.30

17. A car rolls freely down a hill with an increasing speed.
    (a) Draw a diagram to show all of the forces acting on the car.
    (b) What is the direction of the net force on the car?
    (c) What is the largest single force acting on the car?
    (d) When the car reaches a horizontal surface it slows, eventually coming to a stop. Why does this happen?
18. A cyclist of mass 60 kg is riding at a constant speed up a hill that is inclined at 30° to the horizontal. The mass of the bicycle is 20 kg. Figure 10.31 shows the forces acting on the bicycle–cyclist system.

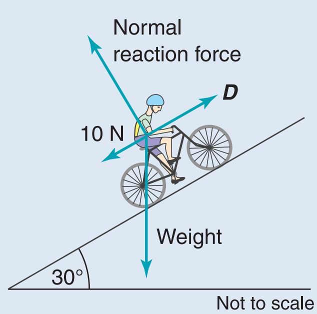


Figure 10.31

    (a) Calculate the net force on the bicycle–cyclist system.
    (b) The sum of the magnitudes of the road friction and air resistance on the system is 10 N. What is the magnitude of the component of the weight of the system that is parallel to the road surface?
    (c) Calculate the magnitude of the driving force ***D***.
    (d) Calculate the magnitude of the normal reaction force on the bicycle–cyclist system.
19. An experienced downhill skier with a mass of 60 kg (including skis) is moving with increasing speed down a slope inclined at 30°. She is moving in a straight line down the slope.
    (a) Calculate the direction of the net force on the skier.
    (b) Draw a diagram showing the forces acting on the skier. Show all of the forces as if they were acting through her centre of mass.
    (c) Calculate the magnitude of the component of the skier's weight that is parallel to the slope.
    (d) If the sum of the forces resisting the movement of the skier down the slope is 8.0 N, calculate the magnitude of the net force on her.
20. A ball of mass 0.50 kg is thrown vertically upwards.
    (a) Calculate the velocity of the ball at the top of its flight.
    (b) Calculate the magnitude of its acceleration at the top of its flight.
    (c) Calculate the net force on the ball at the top of its flight.
21. Calculate the magnitude of the net force on each of the following objects:
    (a) a 1600 kg car while it is accelerating from 0 to 72 km h^{-1} (20 m s^{-1}) in 5.0 s
    (b) a 500 tonne Manly ferry while it is cruising at a constant speed of 20 km h^{-1}
    (c) a space shuttle at lift-off, when its acceleration is 3.0 m s^{-2} and its lift-off mass is 2.2×10^6 kg.
22. A car of mass 1200 kg starts from rest on a horizontal road and a forward thrust of 10 000 N is applied. The resistance to motion due to road friction and air resistance totals 2500 N.
    (a) Calculate the magnitude of the net force on the car.
    (b) Calculate the magnitude of the acceleration of the car.
    (c) Calculate the speed of the car after 5.0 s.
    (d) Calculate the distance the car has travelled after 5.0 s.
23. A train of mass 8.0×10^6 kg, travelling at a speed of 30 m s^{-1}, brakes and comes to rest in 25 s with a constant deceleration.
    (a) Calculate the frictional force acting on the train while it is decelerating.
    (b) Calculate the stopping distance of the train.

CHAPTER REVIEW

24. A physics teacher decides, just for fun, to use bathroom scales (calibrated in newtons) in an elevator. The scales provide a measure of the force with which they push up on the teacher. When the lift is stationary the reading on the bathroom scales is 823 N. Calculate the reading on the scales when the elevator is:
(a) moving upwards at a constant speed of 2.0 m s^{-1}
(b) accelerating downwards at 2.0 m s^{-2}
(c) accelerating upwards at 2.0 m s^{-2}.

25. A roller-coaster carriage (and its occupants), with a total mass of 400 kg, rolls freely down a straight part of the track inclined at 40° to the horizontal with a constant acceleration. The frictional force on the carriage is a constant 180 N. Assume that air resistance is insignificant. What is the magnitude of the acceleration of the carriage?

26. A skateboarder with a mass of 56 kg is rolling freely down a straight incline. The motion of the skateboarder is described in the graph in figure 10.32. Assume that air resistance is insignificant.

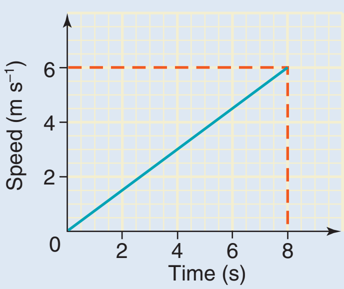


Figure 10.32

(a) Calculate the magnitude of the net force on the skateboarder.
(b) If the friction force resisting the motion of the skateboarder is a constant 140 N, at what angle is the slope inclined to the horizontal?

27. The magnitude of the air resistance, R, on a car can be approximated by the formula:
$R = 1.2\, v^2$
where R is measured in newtons and v is the speed of the car in m s^{-1}.
(a) Design a spreadsheet to calculate the magnitude of the force of air resistance and the net force on the car for a range of speeds as it accelerates from 20 km h^{-1} to 60 km h^{-1} on a horizontal road. Assume that while accelerating, the driving force is a constant 1800 N and the road friction on the non-driving wheels is a constant 300 N.
(b) Use your spreadsheet to plot a graph of the net force versus speed for the car.
(c) Modify your spreadsheet to show how the net force on the car changes when the same acceleration (from 20 km h^{-1} to 60 km h^{-1}) is undertaken while driving up a road inclined at 10° to the horizontal.

28. A 6 kg bowling ball and a 60 kg gold bar are dropped at the same instant from the third floor of the Leaning Tower of Pisa. Use Newton's Second Law of Motion to explain why:
(a) they both reach the ground at the same time
(b) a 6 kg doormat dropped from the same location at the same time takes longer to reach the ground.

29. Copy and complete the following table by fully describing the missing half of the action–reaction pairs.

You push on a wall with the palm of your hand.	
Your foot pushes down on a bicycle pedal.	
The ground pushes up on your feet while you are standing.	
The Earth pulls down on your body.	
You push on a broken-down car to try to get it moving.	
A hammer pushes down on a nail.	

30. What force provides the forward thrust that gets you moving when you are:
(a) cycling
(b) downhill skiing
(c) water skiing
(d) skateboarding
(e) swimming
(f) rowing?

31. Two loaded trolleys of masses 3.0 kg and 4.0 kg (which are joined by a light string) are pulled by a spring balance along a smooth horizontal laboratory bench as shown in figure 10.33. The reading on the spring balance is 14 N.

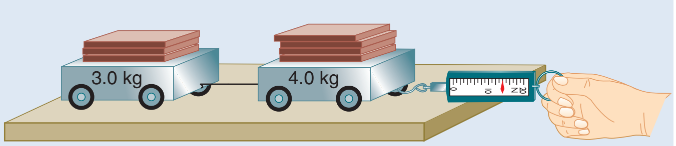


Figure 10.33

(a) Calculate the acceleration of the trolleys.
(b) Calculate the magnitude of the tension in the light string joining the two trolleys.
(c) Calculate the net force on the 4.0 kg trolley.
(d) Calculate the acceleration of the 4.0 kg trolley if the string was cut.

32. A warehouse worker applies a force of 420 N to push two crates across the floor as shown in figure 10.34. The friction force opposing the motion of the crates is a constant 2.0 N for each kilogram.

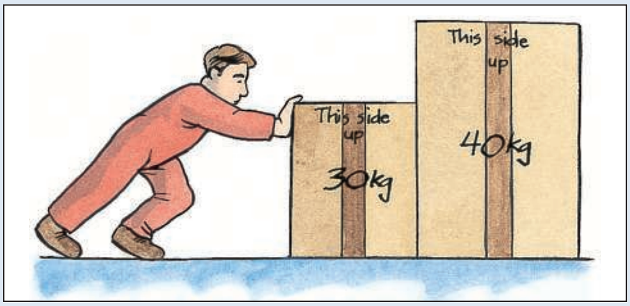


Figure 10.34

(a) Calculate the acceleration of the crates.
(b) Calculate the net force on the 40 kg crate.
(c) Calculate the force exerted by the 40 kg crate on the 30 kg crate.
(d) Calculate the force exerted by the 30 kg crate on the 40 kg crate.
(e) Would the worker find it any easier to give the crates the same acceleration if the positions of the two crates were reversed? Support your answer with calculations.

33. A well-coordinated in-line skater is playing with a yo-yo while accelerating on a horizontal surface. Figure 10.35 shows that when the yo-yo is at its lowest point it makes an angle of 5° with the vertical. Determine the acceleration of the in-line skater.

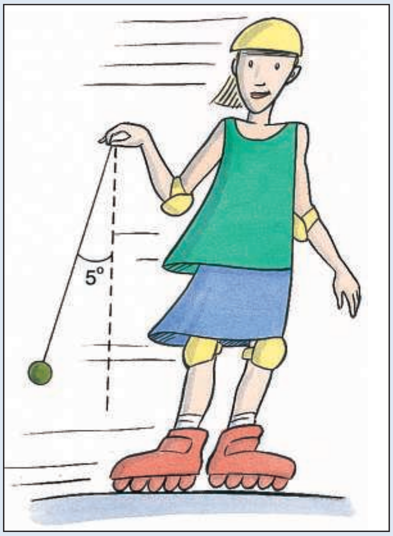


Figure 10.35

34. A car is driven at a constant speed of 8.5 m s^{-1} around a curve with a radius of 6.0 m. The total mass of the car and its passengers is 1500 kg. Calculate:
(a) the acceleration of the car including its passengers
(b) the net force acting on the car including its passengers
(c) the net force acting on the car's 60 kg driver.

35. A racing car of mass 800 kg is driven around a banked circular racetrack at a constant speed of 50 m s^{-1} (180 km h^{-1}). The centripetal acceleration of the racing car is measured to be 4.0 m s^{-2}.
(a) Evaluate the magnitude of the centripetal force acting on the racing car.
(b) What is the magnitude and direction of the net force acting on the racing car?
(c) Calculate the radius of the racetrack.
(d) With the aid of a diagram, explain how the banking of the racetrack allows the racing car to travel faster than it could without banking.

36. A model car set has a circular track that is banked towards the centre of the circle. Two frictional forces can be considered to be acting on a model car as it races around the track. One of them acts at a tangent to the circle in the direction of motion of the car. The other is sideways friction.
(a) Draw labelled diagrams showing all of the forces acting on a model car as it races around the track. You will need to draw one view from the side and one view from the rear.
(b) List the forces that directly influence the magnitude of the centripetal force acting on the model car.

37. An 800 kg racing car travels at constant speed around a slightly banked circular racing track wth a radius of 320 m. The centripetal force on the car is 7840 N.
(a) What is the magnitude and direction of the net force acting on the car?
(b) Calculate the maximum constant speed at which the car can be driven around the circular track.

CHAPTER REVIEW

10.1 FORCE AS A VECTOR

Aim

(a) To show that force is a vector and that the net force is the vector sum of all the forces acting on an object

(b) To analyse the forces acting on an object by resolving the forces into components

Apparatus

three spring balances (5 N)
slotted masses (set of nine 50 g masses and carrier)
marking pen
sheet of A4 paper
masking tape
protractor

Theory

When a point is stationary, the net force acting at that point is zero. We know this because if the point is stationary, it is not changing its motion. The net force is the vector sum of all the forces acting at the point. If the net force at a point is zero, the components of the forces in any direction will add up to zero.

Method

1. Check that the spring balances are 'zeroed' and test them for accuracy by weighing known masses.
2. Using three 5 N spring balances, apply three small forces horizontally to a point, P, so that the point is in equilibrium (see figure 10.36). Use masking tape to secure the ends of the spring balances in place while maintaining the tension so that the net force at the point P is zero. The point P is the point at which the three hooks are in contact.
3. Place a sheet of A4 paper on the table beneath the point P. Use the protractor to measure the angle θ. You need to think carefully about the best way to ensure that the directions are as shown.
4. Draw a diagram of the situation, showing the spring balances and the point P, and label the angles.
5. Draw a separate vector diagram showing the point P and the three forces acting at point P.

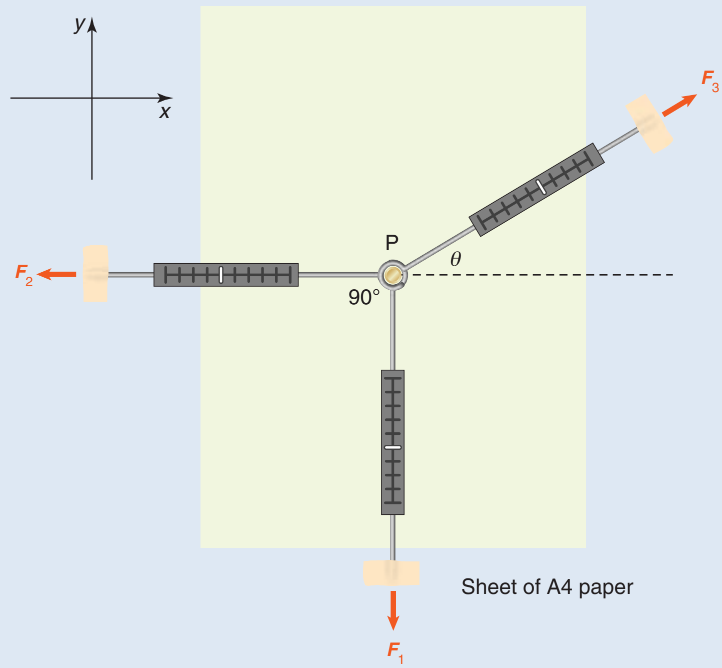


Figure 10.36 *Set up three spring balances horizontally as shown.*

Analysis and questions

Determine the net force acting on point P using the two methods below.

(a) Vector addition method

Apply the 'head to tail' rule for vector addition to all forces. Take care when transferring vectors.

1. Label the net force clearly and state its magnitude and direction.
2. What is the expected magnitude and direction of the net force?
3. Account for any difference between your measured net force and the expected net force.

(b) Component method

1. Transfer your original vector diagram carefully onto graph paper with point P at the origin.
2. Use your graph to find the 'x' component of each of the three forces. Add the 'x' components to obtain the sum of the 'x' components. Repeat the process for the 'y' components.
3. Summarise your results in a table like the one below.

FORCE	'X' COMPONENT (N)	'Y' COMPONENT (N)
F_1		
F_2		
F_3		
Sum		

4. How does the sum of the 'x' components of the three forces compare with the expected value of the sum?
5. How does the sum of the 'y' components of the three forces compare with the expected value of the sum?

10.2 NEWTON'S SECOND LAW OF MOTION

Aim

(a) To examine the relationship between the net force acting on a system, the mass of the system and its acceleration

(b) To use Newton's Second Law of Motion to determine the mass of an object

Apparatus

low-friction trolley
timing and recording device (e.g. ticker-timer, photogates, motion detector and computer interface)
pulley
light string
slotted masses (set of nine 50 g masses and carrier)
metre rule
balance suitable for measuring the mass of the trolley

Theory

Newton's Second Law of Motion describes the relationship between the acceleration of an object, the net force acting on it, and the object's mass. It can be expressed as $\Sigma \boldsymbol{F} = ma$.

Method

1. Use the balance to measure the mass of the trolley. Record its mass.
2. Place 400 g of slotted masses on the trolley. Connect a load of 100 g to the trolley with a light string over a pulley as shown in figure 10.37. The load provides a known external force on the system of the trolley and all of the slotted masses. The magnitude of this external force is equal to the magnitude of the weight of the load.
3. Use your timing and recording device to collect data that will allow you to determine the acceleration of the trolley or glider at several instants as the load is falling.

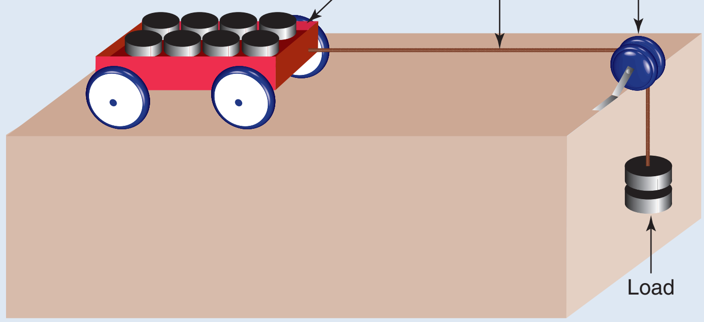


Figure 10.37

4. Repeat this procedure for different loads by taking 100 g from the trolley and adding it to the load. That changes the load, and therefore the external force on the system, without changing the mass of the system. Continue to repeat the procedure until you have removed all of the slotted masses from the trolley.

Analysis

1. Use your data to determine the average acceleration of the system for each external force.
2. Summarise your data in a table which shows the force applied to the system by each external force and the corresponding acceleration of the system.
3. Use your table to plot a graph of external force versus acceleration.
4. Use your graph to make an estimate of the mass of the system of the slotted masses and trolley.

Questions

1. If the force applied by the load through the string was the only horizontal force acting on the trolley, where would the graph cross the vertical axis?
2. What quantity does the intercept on the vertical axis represent?
3. Using your estimate of the mass of the system, what is your estimate of the mass of the trolley?
4. How does your estimate of the mass of the trolley compare with the mass measured by the balance? Suggest reasons for differences between the estimated mass and the measured mass.

CHAPTER 11 MECHANICAL INTERACTIONS

Figure 11.1 *This collision is a mechanical interaction. The motion of a car is changed as a result of the action of a force. The change in motion depends on the size of the force and the mass of the car. However, it is obvious that it's not just the motion of the car that has changed. Some of the car's kinetic energy has been transferred to the object with which it has collided — making it vibrate and even changing its shape. Some of the car's kinetic energy has been transformed into other forms of energy — for example, sound, heat and energy stored in the deformed panels. An understanding of mechanical interactions such as this can teach us how to design safer cars, save countless lives and reduce serious injuries.*

Remember

Before beginning this chapter, you should be able to:

- qualitatively analyse common events involving motion in terms of Newton's laws of motion.

Key content

At the end of this chapter you should be able to:

- recognise that moving objects possess kinetic energy
- define work done on an object and identify its relationship with an increase in energy of that object
- describe energy transformations that take place during collisions
- define the Law of Conservation of Energy
- define momentum as the product of mass and velocity
- define impulse as the product of force and the time interval during which the force acts
- explain why momentum is conserved in collisions if there are no unbalanced external forces acting
- evaluate the benefits of some of the safety features of motor vehicles, including crumple zones, in terms of the concepts of impulse and momentum
- explain conservation of momentum in terms of Newton's Third Law of Motion
- explain the reasoning behind the introduction of low-speed zones in built-up areas.

11.1 THE CONCEPT OF ENERGY

Energy can be defined as the capacity to do work. It is a scalar quantity.

Work is done when an object moves in the direction of a force applied to it. The amount of work done is the product of the magnitude of the force and the displacement of the object in the direction of the force. Work is a scalar quantity.

The word **energy** is often used to describe the way that you feel. For example, you might say 'I don't have a lot of energy today' or on a better day you might say 'I have enough energy to run a marathon'. The word 'energy' is also used to describe something that food has. In each of these cases, the word 'energy' is being used to describe something that provides you with the capacity to make something move. It could be a heavy object, a bicycle or even your own body. Most dictionaries and some physics textbooks define energy as the capacity to do **work**. Work is done when an object moves in the direction of a force applied to it.

The following list of some of the characteristics of energy provides some further clues as to what it really is.

- All matter possesses energy.
- Energy is a scalar quantity — it does not have a direction.
- Energy takes many different forms. It can therefore be classified. Light energy, sound energy, thermal energy, kinetic energy, gravitational potential energy, chemical energy and nuclear energy are some of the different forms of energy.
- Energy can be stored, transferred to other matter or transformed from one form into another. For example, when you hit a cricket ball with a bat, energy is transferred from the bat to the ball. When you dive into a swimming pool, gravitational potential (stored) energy is transformed into kinetic energy.
- Some energy transfers and transformations can be seen, heard, felt, smelt or even tasted.
- It is possible to measure the quantity of energy transferred or transformed.
- Energy cannot be created or destroyed. This statement is known as the Law of Conservation of Energy. The quantity of energy in the universe is a constant. However, nobody knows how much energy there is in the universe.

It is important to distinguish between the words **transferred** *and* **transformed** *when describing energy changes.*

- *The word* **transfer**, *when used as a verb, is defined in* **The Macquarie Dictionary** *as meaning, among other things: '. . . to convey or remove from one place, person, etc. to another'. Energy can be transferred from one object to another.*
- *The word* **transform** *is defined in* **The Macquarie Dictionary** *as meaning, among other things: '. . . to change in form; change to something of a different form; metamorphose . . .'. Energy can be transformed from one form into another form.*

11.2 TRANSFERRING ENERGY

Energy can be transferred to or from matter in several different ways. Energy can be transferred by:

- emission or absorption of electromagnetic or nuclear radiation
- heating and cooling an object or substance as a result of a temperature difference
- the action of a force on an object resulting in movement.

The transfer of energy by the action of a force is called mechanical energy transfer.

Getting down to work

When mechanical energy is transferred to or from an object, the amount of mechanical energy transferred is called work.

The work, W, done when a force, $\boldsymbol{F}$, causes a displacement of magnitude s, in the direction of the force, is defined as:

work = magnitude of the force × displacement in the direction of the force

$$W = F\,s.$$

Work is a scalar quantity. The SI unit of work is the joule. One joule of work is done when a force with a magnitude of one newton causes a displacement of one metre in the same direction as the force. That is, 1 J = 1 N × 1 m = 1 N m. Because energy is a measure of the capacity to do work, the SI unit of energy is also the joule.

Kinetic energy

Kinetic energy is the energy associated with the movement of an object.

Kinetic energy is the energy associated with the movement of an object. By imagining how much energy it would take to make a stationary object move, it can be deduced that kinetic energy depends on the mass and speed of the object.

The change in kinetic energy of an object is equal to the work done on it by the net force acting on it. If an object initially at rest is acted on by a net force of magnitude ΣF and moves a distance s (which will necessarily be in the direction of the net force), its change in kinetic energy, ΔE_k, can be expressed as:

$$\Delta E_k = \Sigma F\, s.$$

The quantity of kinetic energy it possesses is:

$$E_k = \Sigma F\, s$$

because the initial kinetic energy was zero.

Applying Newton's Second Law ($\Sigma \boldsymbol{F} = m\boldsymbol{a}$) to this expression:

$$E_k = mas$$

where

m is the mass of the object and a is its acceleration.

The movement of the object can also be described in terms of its final velocity $\boldsymbol{u}$ and its initial velocity $\boldsymbol{v}$. The magnitudes of the quantities a, s, v and u are related to each other by the equations:

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

and $$s = r = \tfrac{1}{2}(u + v)\ t.$$

Substituting into the expression for kinetic energy:

$$\begin{aligned} E_k &= mas \\ &= m \times \frac{(v - u)}{t} \times \tfrac{1}{2}(u + v)\ t \\ &= \tfrac{1}{2} \times m \times (v^2 - u^2) \\ &= \tfrac{1}{2}mv^2 - \tfrac{1}{2}mu^2. \end{aligned}$$

Because the object was originally at rest, $u = 0$.

The kinetic energy of an object of mass m and speed v can therefore be expressed as:

$$E_k = \tfrac{1}{2}mv^2.$$

SAMPLE PROBLEM 11.1

Doing work to change kinetic energy

A trailer is being pulled along a straight, rough, horizontal road by a car. The trailer and the car travel at a constant speed of 50 km h^{-1}. The forward force applied to the trailer by the car is 4000 N. Frictional forces oppose this force.

(a) In moving a horizontal distance of 500 metres, how much work is done on the trailer by:

(i) the car? (ii) the net force? (iii) the force of gravity?

(b) If the force applied to the trailer by the car is increased to 5000 N and nothing else changes, how much kinetic energy is gained by the trailer over the distance of 500 metres?

SOLUTION

(i)
$$\begin{aligned} W &= F\,s \\ &= 4000 \times 500 \\ &= 4 \times 10^3 \times 5 \times 10^2 \\ &= 2.0 \times 10^6 \text{ J} \end{aligned}$$

The work done on the trailer by the car is 2.0×10^6 J.

(ii) The work done on the trailer by the net force is equal to the change in kinetic energy of the trailer. The trailer is travelling at constant speed, so there is no change in kinetic energy.
No work is done by the net force.

(iii) The work done on the trailer by the force of gravity is zero because the force of gravity has no component in the direction of motion.

(c) When the towing force was 4000 N, the net force was zero. The towing force balanced frictional forces. When the towing force is increased to 5000 N, the net force becomes 1000 N in the direction of motion of the trailer.

$$\begin{aligned} \Delta E_k &= \Sigma F\,s \\ &= 1000 \times 500 \\ &= 500\,000 \text{ J} \end{aligned}$$

The kinetic energy gained is 5.0×10^5 J.

SAMPLE PROBLEM 11.2

Kinetic energy calculations

Compare the kinetic energy of a 100 m Olympic track athlete with that of a family car travelling through the suburbs.

Estimate the mass of the athlete to be 70 kg and the speed of the athlete to be 10 m s^{-1}. Estimate the total mass of the car and its passengers to be 1500 kg and the speed of the car to be about 60 km h^{-1} (17 m s^{-1}).

SOLUTION

For the athlete: $m = 70$ kg, $v = 10$ m s^{-1}

$$\begin{aligned} E_k &= \tfrac{1}{2} mv^2 \\ &= \tfrac{1}{2} \times 70 \times (10)^2 \\ &= 3.5 \times 10^3 \text{ J} \end{aligned}$$

For the car: $m = 1500$ kg, $v = 17$ m s^{-1}

$$\begin{aligned} E_k &= \tfrac{1}{2} mv^2 \\ &= \tfrac{1}{2} \times 1500 \times (17)^2 \\ &= 2.2 \times 10^5 \text{ J} \end{aligned}$$

The value of the ratio $\dfrac{E_k(\text{car})}{E_k(\text{athlete})} = \dfrac{2.2 \times 10^5}{3.5 \times 10^3} = 63.$

The car has about 60 times as much kinetic energy as the athlete.

PHYSICS FACT

The truth of the advertising slogan 'Speed kills' can be appreciated by comparing the kinetic energy of a 1500 kg car travelling at 60 km h^{-1} (16.7 m s^{-1}) with the same car travelling at 120 km h^{-1} (33.3 m s^{-1}).

At 60 km h^{-1} its kinetic energy is:

$$E_k = \tfrac{1}{2}mv^2$$
$$= \tfrac{1}{2} \times 1500 \times (16.7)^2$$
$$= 2.09 \times 10^5 \text{ J}.$$

At 120 km h^{-1} its kinetic energy is:

$$E_k = \tfrac{1}{2}mv^2$$
$$= \tfrac{1}{2} \times 1500 \times (33.3)^2$$
$$= 8.32 \times 10^5 \text{ J}.$$

A doubling of speed produces a four-fold increase in the kinetic energy and, therefore, a four-fold increase in the work that needs to be done on the car to stop it during a crash. It also means that four times as much energy has to be transformed into the energy of deformation, heat and sound or transferred to other objects.

The need for fuel

The kinetic energy of a moving vehicle cannot be created — it must be transferred from another object or transformed from another form of energy. The kinetic energy of the coach in figure 11.2 is transferred from the horses to the coach. But the kinetic energy of the horses is transformed from the chemical energy in the horses' muscles. The kinetic energy of the bicycle in figure 11.2 is transferred from the cyclist to the bicycle. However, the kinetic energy of the cyclist's legs is transformed from chemical energy in the cyclist's muscles. The kinetic energy of a car is transformed from the chemical energy stored in petrol or gas. The kinetic energy of an electric train is transformed from electrical energy. In all of these cases, the kinetic energy is transformed from the chemical energy stored in a fuel. In the case of the horses and cyclist, the fuel is food. In the case of the car it is petrol or gas. The kinetic energy of the electric train might be transformed from a fuel (usually coal in Australia) or from the kinetic energy of water falling from a dam.




Figure 11.2 *Each of these vehicles acquires kinetic energy as the result of a transformation from the chemical energy stored in a fuel.*

11.3 ENERGY TRANSFORMATIONS IN COLLISIONS

When a vehicle collides with a stationary object or another vehicle, some of its kinetic energy is transferred to the object or vehicle. Some of the kinetic energy is still possessed by the vehicle if it does not stop. The remaining kinetic energy is transformed into other forms of energy. These forms include:

Potential energy is the energy stored in an object.

- **potential energy** of deformation. Potential energy of deformation is the energy stored in an object as a result of changing its shape. Sometimes that potential energy of deformation can be easily transformed back into other forms when the object returns to its original shape. For example, the energy of deformation stored in the spring in the jack-in-the-box (figure 11.3) is transformed into kinetic energy when the unsuspecting victim opens the box. When the panels of a car are deformed, it is much more difficult to return them to their original shape.
- sound energy. Sound energy is transmitted through the air as a result of vibrating particles. When a vehicle collides with an object or another vehicle, some of its kinetic energy is transferred to the surrounding air, causing it to vibrate rapidly.
- thermal energy. Thermal energy is energy that a substance possesses as a result of the random motion of the particles within the substance. The vehicle's panels, tyres and other parts get very hot during the collision as kinetic energy is transferred to the particles within them. The other object or vehicle, and even the immediate surrounding road and air, are also heated.

The Law of Conservation of Energy applies to vehicle collisions, as it does to all interactions. The car's kinetic energy is never really 'lost' or destroyed. It is all transferred to other objects or transformed into other forms.

Figure 11.3 *A jack-in-the-box. The potential energy of deformation is transformed into kinetic energy when the spring is allowed to resume its original shape.*

PHYSICS FACT

Riding a bicycle on a horizontal surface requires less energy than running on the same surface at the same speed. Running at a speed of 4.0 m s^{-1} requires a mechanical energy output of about 300 J each second. Riding a bicycle at 4.0 m s^{-1} only requires about 50 J each second. Riding a bicycle on a horizontal surface requires less energy, and is less tiring than running because:

- the whole body of the rider does not rise and fall as it does while running, eliminating the changes in gravitational potential energy (energy stored in an object by raising it)
- where the rider is seated, the muscles need to transform much less chemical energy to support body weight. The strongest muscles in the body can be used almost exclusively to turn the pedals.

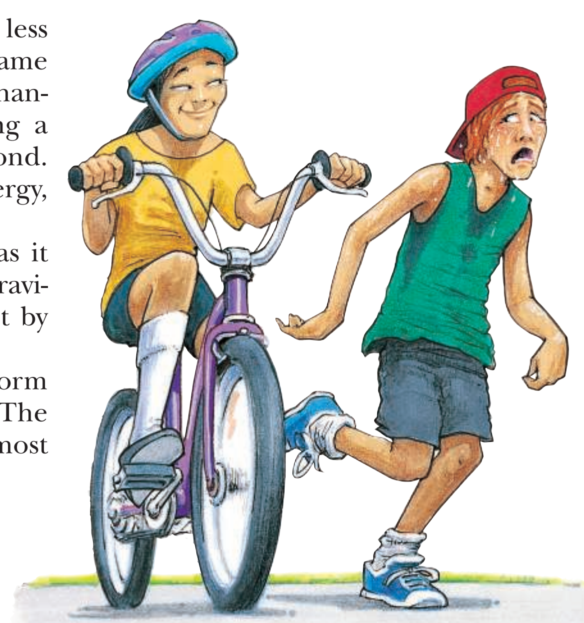


Figure 11.4 *Riding a bicycle on a horizontal surface requires less energy than running at the same speed on the same surface.*

11.4 MOMENTUM

How difficult is it to stop a moving object? How difficult is it to make a stationary object move? The answer to both of these questions depends on two physical characteristics of the object:

- the object's mass
- how fast the object was moving, or how fast you want it to move.

The product of these two physical characteristics is called **momentum**. The momentum, $\boldsymbol{p}$, of an object of mass, m, with a velocity, $\boldsymbol{v}$, is defined as: $\boldsymbol{p} = m\boldsymbol{v}$.

Momentum is the product of the mass of an object and its velocity. It is a vector quantity.

Momentum is a vector quantity and has SI units of kg m s^{-1}.

SAMPLE PROBLEM 11.3

Calculating the momentum of a train

What is the momentum of a train of mass 8.0×10^6 kg that is travelling at a speed of 15 m s^{-1} in a northerly direction?

SOLUTION

$m = 8.0 \times 10^6$ kg, $\boldsymbol{v} = 15$ m s^{-1} north

$$\begin{aligned} \boldsymbol{p} &= m\boldsymbol{v} \\ &= 8.0 \times 10^6 \times 15 \\ &= 1.2 \times 10^7 \text{ kg m s}^{-1} \text{ north} \end{aligned}$$

Making an object stop, or causing it to start moving, requires a non-zero net force. The relationship between the net force applied to an object and its momentum can be explored by applying Newton's Second Law of Motion to the object.

$$\begin{aligned} \Sigma\boldsymbol{F} &= m\boldsymbol{a} \\ \Sigma\boldsymbol{F} &= m\frac{\Delta\boldsymbol{v}}{\Delta t} \\ \Sigma\boldsymbol{F}\Delta t &= m\Delta\boldsymbol{v} \end{aligned}$$

Impulse is the product of the force and the time interval over which it acts. Impulse is a vector quantity with SI units of N s.

The product $\Sigma\boldsymbol{F}\Delta t$ is called the **impulse** of the net force. The impulse of any force is defined as the product of the force and the time interval over which it acts. Impulse is a vector quantity with SI units of N s.

The product $m\Delta\boldsymbol{v}$ is the change in momentum.

$$\begin{aligned} m\Delta\boldsymbol{v} &= m(\boldsymbol{v} - \boldsymbol{u}) \\ &= m\boldsymbol{v} - m\boldsymbol{u} \\ &= \boldsymbol{p}_\text{f} - \boldsymbol{p}_\text{i} \end{aligned}$$

where

$\boldsymbol{p}_\text{f}$ = final momentum of the object

$\boldsymbol{p}_\text{i}$ = initial momentum of the object.

PRACTICAL ACTIVITIES

11.1

Impulse and change in momentum

Thus, the effect of a net force on the motion of an object can be summarised by the statement: impulse = change in momentum.

In fact, Newton's Second Law of Motion, when translated from the original Latin reads:

The rate of change of momentum is directly proportional to the magnitude of the net force and is in the direction of the net force.

It can be expressed as:

$$\Sigma\boldsymbol{F} = \frac{\Delta\boldsymbol{p}}{\Delta t}.$$

SAMPLE PROBLEM 11.4

Momentum and impulse of a car

A 1200 kg car collides with a concrete wall at a speed of 15 m s^{-1} and takes 0.06 s to come to rest.

(a) What is the change in momentum of the car?

(b) What is the impulse on the car?

(c) What is the magnitude of the force exerted by the wall on the car?

(d) What would be the magnitude of the force exerted by the wall on the car if the car bounced back from the wall with a speed of 3.0 m s^{-1} after being in contact for 0.06 s?

SOLUTION

(a) Assign the initial direction of the car as positive.

$m = 1200$ kg, $u = 15$ m s^{-1}, $v = -3.0$ m s^{-1}, $\Delta t = 0.06$ s

$$\begin{aligned}\Delta \boldsymbol{p} &= m\boldsymbol{v} - m\boldsymbol{u} \\ &= m(\boldsymbol{v} - \boldsymbol{u}) \\ &= 1200\ (0 - 15) \\ &= 1200 \times -15 \\ &= -1.8 \times 10^4 \text{ kg m s}^{-1}\end{aligned}$$

The change in momentum is 1.8×10^4 kg m s^{-1} in a direction opposite to the original direction of the car.

(b) Impulse on the car = change in momentum of the car

$$= -1.8 \times 10^4 \text{ kg m s}^{-1}$$

The impulse on the car is 1.8×10^4 N s in a direction opposite to the original direction of the car.

(c) Magnitude of impulse = $F\Delta t$

$$\begin{aligned}1.8 \times 10^4 &= F \times 0.06 \\ F &= \frac{1.8 \times 10^4}{0.06} \\ &= 3.0 \times 10^5 \text{ N}\end{aligned}$$

(d) Impulse = $m\Delta v$

$$\begin{aligned}&= 1200\ (-3 - 15) \\ &= 1200 \times -18 \\ &= -2.16 \times 10^4 \text{ N s}\end{aligned}$$

$$\begin{aligned}2.16 \times 10^4 &= F\Delta t \\ 2.16 \times 10^4 &= F \times 0.06 \\ F &= \frac{2.16 \times 10^4}{0.06} \\ &= 3.6 \times 10^5 \text{ N}\end{aligned}$$

Impulse from a graph

The force that was determined in sample problem 11.4 was actually the average force on the car. In fact, the force acting on the car is not constant. The impulse delivered by a changing force is given by impulse = $\boldsymbol{F}_{av}\Delta t$. If a graph of force versus time is plotted, the impulse can be determined from the area under graph.

SAMPLE PROBLEM **11.5**

Speed of a roller skater

The graph in figure 11.5 describes the changing horizontal force on a 40 kg roller skater as she begins to move from rest. Estimate her speed after 2.0 seconds.

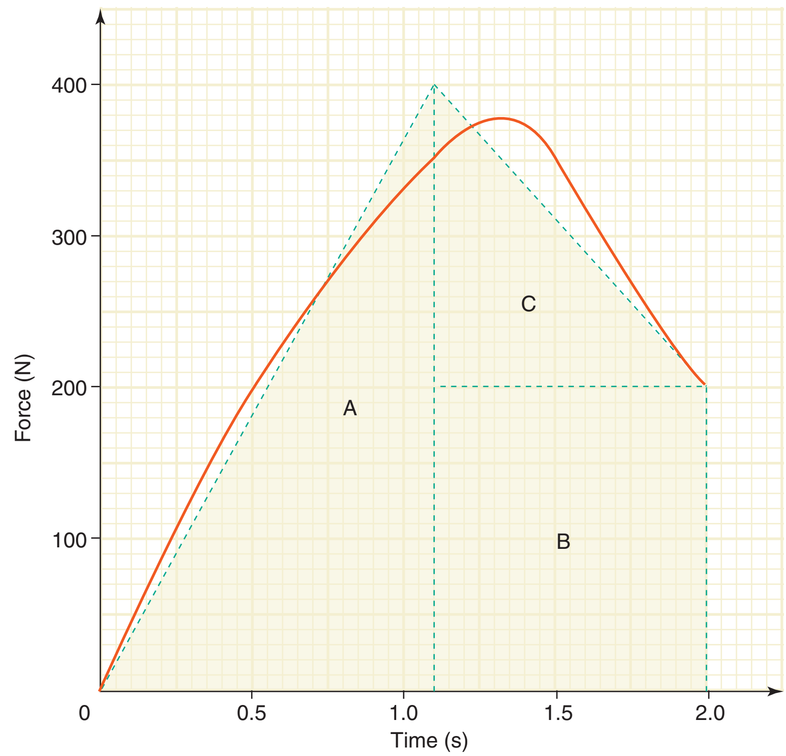


Figure 11.5

SOLUTION

The magnitude of the impulse on the skater can be determined by calculating the area under the graph. This can be determined by either counting squares or by finding the shaded area.

$$\begin{aligned}\text{Magnitude of impulse} &= \text{area A} + \text{area B} + \text{area C}\\ &= \tfrac{1}{2} \times 1.1 \times 400 + 0.9 \times 200 + \tfrac{1}{2} \times 0.9 \times 200\\ &= 220 + 180 + 90\\ &= 490\ \text{N s}\end{aligned}$$

$$\begin{aligned}\text{Magnitude of change in momentum} &= m\Delta v\\ 490 &= 40 \times \Delta v\\ \Delta v &= \frac{490}{40}\\ &= 12\ \text{m s}^{-1}\end{aligned}$$

As her initial speed is zero (she started from rest), her speed after 2.0 seconds is 12 m s^{-1}.

Protecting that frail human body

The human body does not cope very well with sudden blows. The skeleton provides a fairly rigid frame that protects the vital organs inside and, with the help of your muscles, enables you to move.

Figure 11.6 ***A Sydney Kings basketballer is about to complete a 'slam dunk'. When his feet hit the ground, he bends his knees to minimise the possibility of knee injuries, while the impact on his feet will be cushioned by specialised footwear and the springiness of the court's surface.***

A sudden impact to your body, or part of your body, can:

- push or pull the bones hard enough to break them
- tear or strain the ligaments that hold the bones together
- tear or strain muscles or the tendons that join the muscles to bones
- push bones into vital organs like the brain and lungs
- tear, puncture or crush vital organs like the kidneys, liver and spleen.

The damage that is done to you depends on the magnitude of the net force, and subsequent acceleration, to which your body is subjected. In any collision, the net force acting on your body, or part of your body, can be expressed as:

$$\Sigma \boldsymbol{F} = \frac{\Delta \boldsymbol{p}}{\Delta t}.$$

The symbol $\Delta \boldsymbol{p}$ represents the change in momentum of your body or the part of your body directly affected by that net force.

Protecting yourself in sport

When you land on a basketball court after a high leap, the size of the force on your knees depends on the magnitude of your change in momentum and the time interval over which you stop. The change in momentum is beyond your control. The speed of your body when you land, and therefore your change in momentum, is determined by the height from which you drop. However, you do have control over the time interval during which the momentum changes. If Δt can be increased, the magnitude of the net force applied to you will be decreased. In sport, you can do this by:

- bending your knees when you land after jumping in sports such as netball and basketball. This increases the time interval over which your knees change their momentum, and decreases the likelihood of ligament damage.
- moving your hand back when you catch a fast-moving ball in sports such as cricket. The ball changes its momentum over a longer time interval, reducing the force applied to it by your hand. In turn, the equal and opposite reaction force on your hand is less.
- wearing gloves and padding in sports such as baseball, softball and cricket. Thick gloves are essential for wicket-keepers who catch the solid cricket ball while it is travelling at speeds up to 150 km h^{-1}. The gloves decrease the rate of change of momentum of the ball, and consequently the force applied to your hand, by increasing the time taken for the ball to stop.
- wearing footwear that increases the time interval during which your feet stop when they hit the ground. This is particularly important for people who run on footpaths and other hard surfaces. Indoor basketball and netball courts have floors that, although hard, bend a little, increasing the period of impact of running feet.

Protecting yourself on the road

In the event of a car collision, the net force applied to your body as its motion suddenly changes can be controlled in two ways:

1. By reducing your initial momentum and therefore, your change in momentum, by driving at a moderate speed. Of course, by driving at a moderate speed, you are less likely to have a collision in the first place. Low-speed zones, speed humps and strict enforcement of speed limits contribute to making accidents less likely and to reducing injuries when accidents do occur.
2. By increasing the time interval during which the change in momentum of the car, and the change in momentum of its occupants, takes place.

Cars that crumple

eBook*plus*

Weblink:
Car safety systems

Cars are designed to crumple at the front and rear. This provision increases the time interval during which the momentum of the car changes in a collision, further protecting its occupants from death or serious injury. Even though the front and rear of the car crumple, the passenger compartment is surrounded by a rigid frame. The engine is also surrounded by rigid structures that prevent it from being pushed into the passenger compartment. The tendency of the roof to crush is currently being reduced by increasing the thickness of the windscreen and side windows, using stronger adhesives and strengthening the roof panel.

The inside of the passenger compartment is also designed to protect occupants. Padded dashboards, collapsible steering wheels and airbags are designed to reduce the rate of change of momentum of occupants in a collision. Interior fittings like switches, door knobs and the handbrake are sunk so that the occupants do not collide with them.

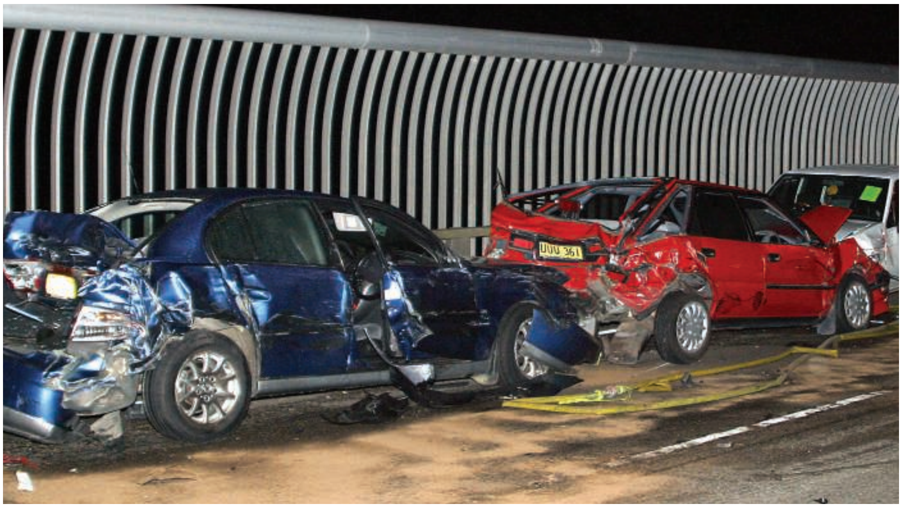


Figure 11.7 *Crumple zones at the front and rear of cars reduce the rate of change of momentum of the car and its occupants during a collision.*

Don’t be an egghead

In a serious bicycle accident, the head is likely to collide at high speed with the road or another vehicle. Even a simple fall from a bike can result in a collision of the head with the road at a speed of about 20 km h^{-1}. Without the protection of a helmet, concussion is likely as the skull decelerates very quickly due to the large net force on it. It will come to rest while the brain is still in motion. The brain will collide with the skull. If the net force on the skull and its subsequent deceleration is large enough, the brain can be severely bruised or torn, often resulting in permanent brain damage or death. The effect is not unlike that of dropping a soft-boiled egg onto a hard floor.

Bicycle helmets typically consist of an expanded polystyrene liner about two centimetres thick, covered in a thin, hard, polymer shell. They are designed to crush on impact. Although a helmet does not guarantee survival in a serious bicycle accident, it does reduce the net force applied to the skull, and therefore increases the chances of survival dramatically. The polystyrene liner of the helmet increases the time interval during which the skull changes its momentum.

Helmets used by motor cyclists, in horse riding, motor racing, cricket and in many other sports serve the same purpose — that is, to increase the time interval over which a change in momentum takes place.

Figure 11.8 ***Helmets save lives and prevent serious injury in many sports. They increase the time interval over which a change in momentum takes place.***

Seatbelts and safety

In a high-speed head-on car collision, each car comes to a stop rapidly. An occupant not wearing a seatbelt continues at the original speed of the car (as described by Newton's First Law of Motion) until acted on by a non-zero net force. An unrestrained occupant therefore moves at high speed until:

- colliding with part of the interior of the car, stopping even more rapidly than the car itself, usually over a distance of only several centimetres (Most deaths and injuries in car crashes are caused by collisions between the occupants and the interior of the car.)
- crashing through the stationary or almost stationary windscreen into the other car or onto the road
- crashing into another occupant closer to the front of the car.

An occupant properly restrained with a seatbelt stops with the car. In a typical suburban crash, the deceleration takes place over a distance of about 50 cm. The rate of change of the momentum of a restrained occupant is much less than that of an unrestrained occupant. Therefore, the net force on a restrained occupant is less. As well as increasing the time interval over which its wearer comes to a stop, a properly fitted seatbelt spreads the force over a larger area of the body.

Inertia-reel seatbelts

Inertia-reel seatbelts allow car occupants some freedom of movement while they are worn. However, in the event of a sudden change in velocity of the car, they lock and restrain the occupant (see figure 11.9). Inertia-reel seatbelts are designed with Newton's First Law of Motion in mind. When the car stops suddenly, a pendulum continues to move forward. Part of the pendulum prevents the reel holding the belt from turning. This locks the belt into place. The name 'inertia reel' is given to these seatbelts because the inertia of the pendulum causes the belt to be locked. Another type of seatbelt uses an electronic sensor. When the sensor detects an unusually large deceleration it releases a gas propellant which causes the reel to be locked.

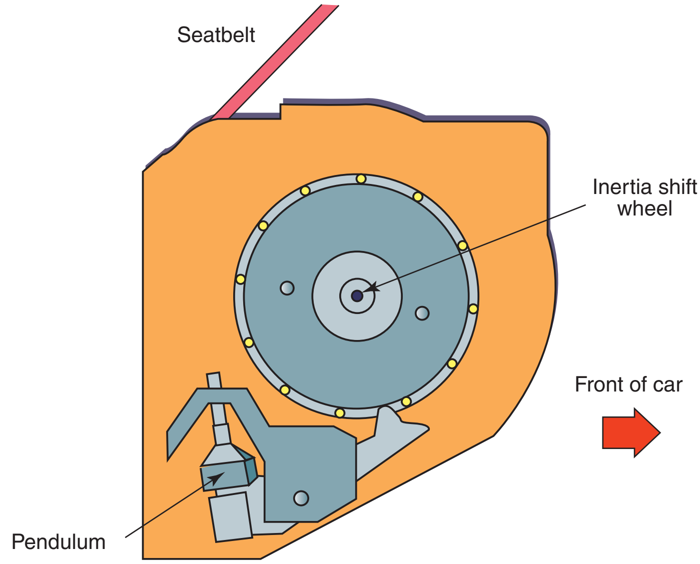


Figure 11.9 *Operation of an inertia-reel seatbelt. This reel is shown in the locked position.*

PHYSICS IN FOCUS

Airbag technology

Airbags are designed to increase the time interval during which the occupant's momentum decreases in a collision, reducing the net force on the occupant. Airbags inflate when the crash sensors in the car detect a large deceleration. When the sensors are activated, an electric current is used to ignite the chemical compound sodium azide (NaN_3), which is stored in a metal container at the opening of the airbag. The sodium azide burns rapidly, producing other sodium compounds and nitrogen gas. The reaction is explosive, causing a noise like the sound of gunfire. The nitrogen gas inflates the airbag to a volume of about 45 litres in only 30 milliseconds.

When the occupant's body makes contact with the airbag, nitrogen gas escapes through vents in the bag. The dust produced when an airbag is activated is a mixture of the talcum powder used to lubricate the bags and the sodium compounds produced by the chemical reaction. Deflation must be rapid enough to allow a driver to see after the accident.

(continued)

Figure 11.10 *Airbags increase the time interval during which the occupant's momentum decreases.*

John Kirkpatrick is a senior project engineer at Holden Innovation in Port Melbourne. After completing a degree in mechanical engineering, John began work with a company that supplied parts to the automotive industry. Some years later, he accepted a job at Holden and has since worked in a variety of vehicle design areas, including the development of vehicle body structures, airbags and other features that maximise the protection offered to occupants in the event of an accident.

Before any physical testing of a new model takes place, the vehicle structure is modelled on a computer to ensure that it has adequate durability, comfort (in terms of noise and vibration for example) and accident performance. The computer modelling is then verified with the first physical testing of real vehicles. Following this, the design will progress through a number of refinements before the new model is ready for sale to the public.

One interesting aspect in the development of an airbag system is the calibration of the sensor which triggers the airbags. Current 'state of the art' technology for driver and passenger airbags uses a single sensing module mounted within the passenger compartment of the vehicle. This module continually monitors the longitudinal acceleration of the car. Complex calculations and comparisons are performed by a micro-processor within the sensing module before it 'decides' whether or not to trigger the airbags.

Many cars are crashed on the computer and in real life during the development of the vehicle structure and airbag system. The crash events used to develop an airbag calibration include high- and low-speed collisions, full and angled frontal impacts and pole- or tree-type collisions.

And what does John like most about his work? Being paid to crash cars, of course!

Figure 11.11 ***Holden safety engineer John Kirkpatrick retrieves collision information from the airbag computer system.***

A closer look at collisions

PRACTICAL ACTIVITIES

11.2

Simulating a collision

eBook*plus*

Weblink:
Elastic and inelastic collisions applet

When two objects collide with each other, each object exerts a force on the other. Newton's Second Law of Motion can be applied to the system of two objects just as it can be applied to each object. By applying the formula $\Sigma \boldsymbol{F} = \frac{\Delta \boldsymbol{p}}{\Delta t}$ to a system of one or more objects, another expression of Newton's Second Law of Motion can be written:

If the net force acting on a system is zero, the total momentum of the system does not change.

This expression is a statement of the Law of Conservation of Momentum. It is often expressed as follows:

If there are no external forces acting on a system, the total momentum remains constant. In symbols, if $\Sigma \boldsymbol{F} = 0$, $\Delta \boldsymbol{p} = 0$.

The two blocks labelled A and B in figure 11.12 comprise a system. The net force on the system is zero. Therefore, the total momentum of the system remains constant.

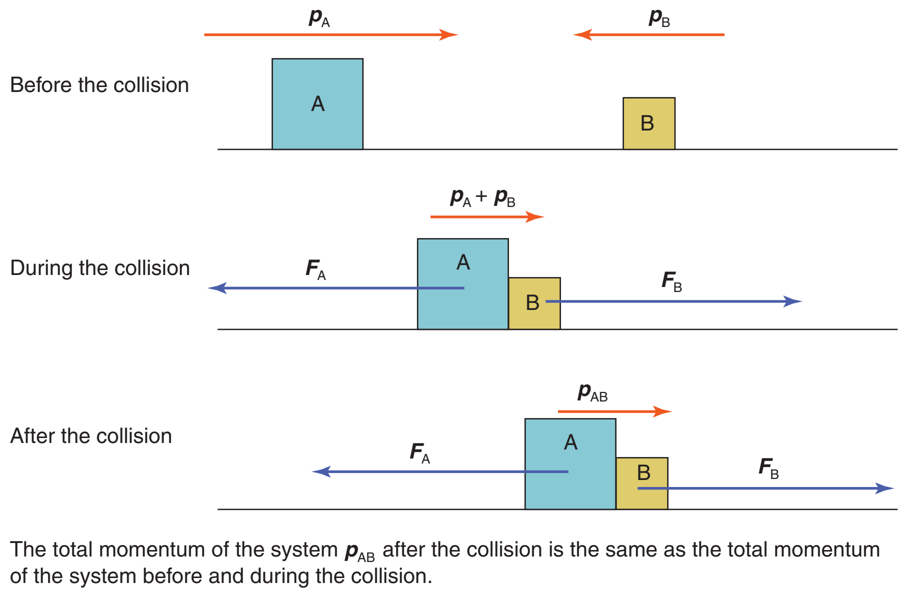


Figure 11.12 *The net force on this system of two blocks is zero.*

SAMPLE PROBLEM 11.6

Momentum of two cars

A 1500 kg car travelling at 12 m s^{-1} on an icy road collides with a 1200 kg car travelling at the same speed, but in the opposite direction. The cars lock together after impact.

(a) What is the momentum of each car before the collision?
(b) What is the total momentum before the collision?
(c) What is the total momentum after the collision?
(d) With what speed is the tangled wreck moving *immediately* after the collision?

SOLUTION

(a) Assign the direction in which the first car is moving as positive.

1500 kg car: $m = 1500$ kg, $v = 12$ m s^{-1}

$$\begin{aligned} p &= mv \\ &= 1500 \times 12 \\ &= 18\,000 \text{ kg m s}^{-1} \end{aligned}$$

1200 kg car: $m = 1200$ kg, $v = -12$ m s^{-1}

$$\begin{aligned} p &= mv \\ &= 1200 \times -12 \\ &= -14\,400 \text{ kg m s}^{-1} \\ p_i &= 18\,000 - 14\,400 \\ &= 3600 \text{ kg m s}^{-1} \end{aligned}$$

(b) The description of the road suggests that friction is insignificant. It can be assumed that there are no external forces acting on the system.

$$\begin{aligned} p_f &= p_i \\ &= 3600 \text{ kg m s}^{-1} \end{aligned}$$

(c) The tangled wreck can be considered as a single mass of 2700 kg.

$m = 2700$ kg, $p_f = 3600$ kg m s^{-1}

$$\begin{aligned} p_f &= mv \\ 2700\,v &= 3600 \\ v &= 1.3 \text{ m s}^{-1} \text{ in the direction of the initial velocity of the first car} \end{aligned}$$

11.5 MOMENTUM AND NEWTON'S THIRD LAW OF MOTION

Newton's Third Law of Motion was discussed in chapter 10 (pages 211–12).

When two cars collide on a smooth surface, the momentum of each car changes because, during the collision, each car has a non-zero net force acting on it. However, the total change in momentum of the two-car system is zero. The total change in momentum, $\Delta \boldsymbol{p}$, can be expressed in terms of the changes in momentum, $\Delta \boldsymbol{p}_1$ and $\Delta \boldsymbol{p}_2$, of each of the two cars.

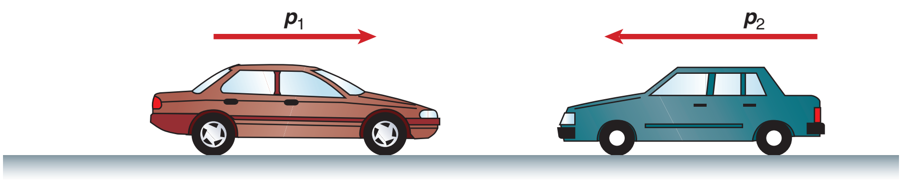


Figure 11.13 *Two cars about to collide on a smooth surface. The total momentum after the collision will be the same as the total momentum before the collision.*

$$\begin{aligned} \Delta \boldsymbol{p} &= \Delta \boldsymbol{p}_1 + \Delta \boldsymbol{p}_2 \\ \Delta \boldsymbol{p}_1 + \Delta \boldsymbol{p}_2 &= 0 \\ \Delta \boldsymbol{p}_1 &= - \Delta \boldsymbol{p}_2 \end{aligned}$$

The change in momentum of the first car is equal and opposite to the change in momentum of the second car.

The change in momentum of the first car is given by:

$$\Delta \boldsymbol{p}_1 = \boldsymbol{F}_1 \Delta t$$

where

$\boldsymbol{F}_1$ = average net force on the first car over the time interval, Δt, during which the collision takes place.

The change in momentum of the second car is given by:

$$\Delta \boldsymbol{p}_2 = \boldsymbol{F}_2 \Delta t$$

where

$\boldsymbol{F}_2$ = average net force on the second car over the time interval, Δt, during which the collision takes place.

The forces acting on each of the cars during the collision can now be compared.

Since
$$\begin{aligned} \Delta \boldsymbol{p}_1 &= - \Delta \boldsymbol{p}_2 \\ \boldsymbol{F}_1 \Delta \mathrm{t} &= - \boldsymbol{F}_2 \Delta \mathrm{t} \\ \boldsymbol{F}_1 &= - \boldsymbol{F}_2 \end{aligned}$$

$\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are equal and opposite in direction. This result is, not surprisingly, totally consistent with Newton's Third Law of Motion.

The interaction between the two cars can be summarised as follows:

- the total momentum of the system of the two cars remains constant
- the total change in momentum is zero
- the change in momentum of the first car is equal and opposite to the change in momentum of the second car
- the force that the first car exerts on the second car is equal and opposite to the force that the second car exerts on the first car.

SAMPLE PROBLEM 11.7

Momentum calculations

A 2000 kg delivery van travelling at 30 m s^{-1} on a smooth surface collides with a small stationary car of mass 1000 kg. The two vehicles lock together and the tangled wreck continues to move in the direction in which the van was travelling.

(a) What is the speed of the tangled wreck immediately after the collision?
(b) What is the change in momentum of the car?
(c) What is the change in momentum of the van?

SOLUTION

(a) Assign the direction in which the van is initially moving as positive.

Van: $m = 2000$ kg, $\boldsymbol{u} = 30$ m s^{-1}

Car: $m = 1000$ kg, $\boldsymbol{u} = 0$

The initial momentum, $\boldsymbol{p}_i$, of the system is given by:

$$\begin{aligned} \boldsymbol{p}_{\text{van}} + \boldsymbol{p}_{\text{car}} &= 2000 \times 30 + 0 \\ &= 60\,000 \text{ kg m s}^{-1} \end{aligned}$$

The momentum of the system after the collision, $\boldsymbol{p}_f$, is the momentum of just one object: the tangled wreck.

$$\boldsymbol{p}_f = 3000 \times \boldsymbol{v}$$

where $\boldsymbol{v}$ is the velocity of the tangled wreck after the collision

But $\boldsymbol{p}_f = \boldsymbol{p}_i$

$$\begin{aligned} 3000\,\boldsymbol{v} &= 60\,000 \\ \boldsymbol{v} &= \frac{60\,000}{3000} \\ &= 20 \text{ m s}^{-1} \end{aligned}$$

The speed of the tangled wreck after the collision is 20 m s^{-1}.

(b)
$$\begin{aligned} \Delta\boldsymbol{p}_{\text{car}} &= m_{\text{car}}\Delta\boldsymbol{v}_{\text{car}} \\ &= 1000 \times (20 - 0) \\ &= 20\,000 \text{ kg m s}^{-1} \text{ in the direction of the van's initial velocity} \end{aligned}$$

(c)
$$\begin{aligned} \Delta\boldsymbol{p}_{\text{van}} &= m_{\text{van}}\Delta\boldsymbol{v}_{\text{van}} \\ &= 2000(20 - 30) \\ &= -20\,000 \text{ kg m s}^{-1} \\ &= 20\,000 \text{ kg m s}^{-1} \text{ opposite to the direction of the van's initial velocity} \end{aligned}$$

SUMMARY

- All moving objects possess kinetic energy. The kinetic energy of an object can be expressed as $E_k = \frac{1}{2}mv^2$.
- The work done on an object can increase its kinetic energy.
- The Law of Conservation of Energy states that energy cannot be created or destroyed. However, energy can be transferred from one object to another or transformed from one form of energy into another.
- The Law of Conservation of Energy applies to vehicle collisions as it does to all interactions.
- When a vehicle collides with a stationary object or another vehicle, some of its kinetic energy can be transferred to the object or vehicle. Some of the kinetic energy is still possessed by the vehicle if it does not stop. The remaining kinetic energy is transformed into other forms of energy.
- The momentum of an object is the product of its mass and its velocity.
- The impulse delivered to an object by a force is the product of the force and the time interval during which the force acts on the object.
- The impulse delivered by the net force on an object is equal to the change in momentum of the object. In symbols, $\Sigma \boldsymbol{F}\Delta t = m\Delta \boldsymbol{v}$.
- The impulse delivered by a force can be found by determining the area under a graph of the force versus time.
- The damage done to the human body during a collision depends on the magnitude of the net force, and subsequent acceleration that it is subjected to.
- The net force on a human body during a collision can be decreased by increasing the time interval during which its momentum changes. Vehicle safety features such as crumple zones, together with seatbelts and airbags, are designed to increase this time interval. Low-speed zones and speed humps encourage people to drive at lower speeds and, therefore, with less momentum — reducing the likelihood of injury when a collision does occur.
- If the net force acting on a system is zero, the total momentum of the system does not change. This statement is an expression of the Law of Conservation of Momentum.
- When two objects collide, the force applied by the first object on the second is equal and opposite to the force applied by the second object on the first.

QUESTIONS

Assume that the magnitude of the gravitational field strength at the Earth's surface is 9.8 N kg^{-1}.

1. How are mechanical energy transfers different from other types of energy transfer?
2. Distinguish between an energy transformation and an energy transfer.
3. How much work is done on a 4.0 kg brick as it is lifted through a vertical distance of 1.5 m?
4. Imagine that you are trying to single-handedly push start a 2000 kg truck with its handbrake on. Not surprisingly, the truck doesn't move. How much work are you doing on the truck?
5. Estimate the kinetic energy of:
   (a) a car travelling at 60 km h^{-1} (16.7 m s^{-1}) on a suburban street
   (b) a tennis ball as it is returned to the server in a Wimbledon final
   (c) a cyclist riding to work
   (d) a snail crawling across a footpath.
6. A car of mass 1200 kg is being towed by a thick rope connected to a larger car. After stopping at traffic lights, the tension in the rope is a constant 4000 N for a distance of 50 metres. The frictional force resisting the motion of the smaller car is 400 N.
   (a) Calculate how much work is done on the smaller car by the net force.
   (b) Evaluate the kinetic energy of the smaller car after the distance of 50 metres has been covered.
   (c) Calculate the speed of the smaller car at the end of the 50 metres of towing.
7. A parked delivery van with a mass of 1800 kg rolls down a hill because its handbrake is accidentally left off. After rolling 50 metres it reaches a speed of 16 m s^{-1}. Calculate the net force acting on the van while it was rolling down the hill.
8. Describe the energy transformations that take place when a speeding car collides with a large, reinforced concrete structure.
9. Make an estimate, to one significant figure, of the magnitude of each of the following:
   (a) the momentum of an Olympic athlete in the 100 m sprint
   (b) the momentum of a family car travelling at the speed limit along a suburban street
   (c) the impulse that causes a 70 kg football player running at top speed to stop abruptly as he collides with an unseen goalpost

(d) the impulse applied to a netball by a goal shooter as she pushes it up towards the goal at a speed of 5 m s^{-1}
(e) the change in momentum of a tennis ball as it is returned to the server in a Wimbledon final.

10. A railway cart of mass 500 kg travelling at 3.0 m s^{-1} due west comes to rest in 2.0 s when the engine pulling it stops.
(a) Calculate the impulse that has been applied to the cart.
(b) Calculate the change in momentum of the cart.
(c) Calculate the magnitude of the average force acting on the cart as it comes to a stop.

11. The graph in the figure 11.14 shows how the net force on an object of mass 2.5 kg changes with time.

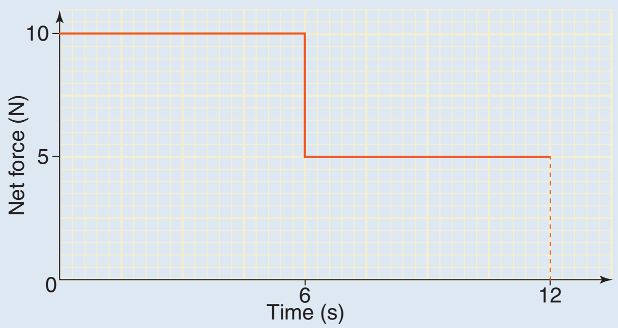


Figure 11.14

(a) Calculate the impulse applied to the object during the first 6.0 s.
(b) If the object was initially at rest, what is its momentum after 12 s?
(c) Draw a graph of acceleration versus time for the object.

12. A car with a total mass of 1400 kg (including occupants), travelling at 60 km h^{-1}, hits a large tree and stops in 0.080 s.
(a) Calculate the impulse that is applied to the car by the tree.
(b) Calculate the force exerted by the tree on the car.
(c) Calculate the magnitude of the deceleration of the 70 kg driver of the car if he is wearing a properly fitted seatbelt.

13. Figure 11.15 shows how the horizontal force on the upper body of each of two occupants of a car changes as a result of a head-on collision. One occupant is wearing a seatbelt while the other is not. Both occupants are stationary 0.10 s after the initial impact.

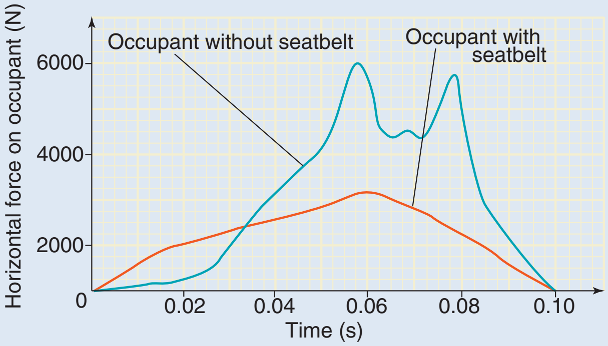


Figure 11.15

(a) What is the horizontal impulse on the occupant wearing the seatbelt?
(b) If the mass of the occupant wearing the seatbelt is 60 kg, determine the speed of the car just before the initial impact.
(c) Is the occupant not wearing the seatbelt heavier or lighter than the other (more sensible) occupant? Write a paragraph explaining the difference in shape between the two curves on the graph.

14. Joggers are advised to run on grass or other soft surfaces rather than concrete paths or bitumen roads to reduce the risk of knee and other leg injuries. Explain why this is so.

15. A 75 kg basketballer lands vertically on the court with a speed of 3.2 m s^{-1}.
(a) What total impulse is applied to the basketballer's feet by the ground?
(b) If the basketballer's speed changes from 3.2 m s^{-1} to zero in 0.10 s, what total force does the ground apply to his feet?

16. Use the ideas presented in this chapter to explain why:
(a) dashboards of cars are padded
(b) cars are deliberately designed to crumple at the front and rear
(c) the compulsory wearing of bicycle helmets has dramatically reduced the number of serious head injuries in bicycle accidents.
A single answer (rather than three separate answers) is acceptable.

17. It is often said that seatbelts prevent passengers from being thrown forwards in a car collision. What is wrong with such a statement?

18. Airbags are fitted to the centre of the steering wheel of many new cars. In the event of a sudden deceleration, the airbag inflates rapidly, providing extra protection for a driver restrained by a seatbelt. Explain how airbags reduce the likelihood of serious injury or death.

19. A toy car with a mass of 2.0 kg collides with a wall at a speed of 1.0 m s^{-1} and rebounds in the opposite direction with a speed of 0.50 m s^{-1}.
   (a) What is the change in momentum of the toy car?
   (b) What is the impulse applied by the toy car to the wall? Explain how you obtained your answer without any information about the change in momentum of the wall.
   (c) Does the wall actually move as a result of the impulse applied by the toy car? Explain your answer.

20. A physics student is experimenting with a low-friction cart on a smooth horizontal surface. Predict the final velocity of the 2.0 kg cart in each of these two experiments.
   (a) The cart is travelling at a constant speed of 0.60 m s^{-1}. A suspended 2.0 kg mass is dropped onto it as it passes.
   (b) The cart is loaded with 2.0 kg of sand. As the cart moves with an initial speed of 0.60 m s^{-1} the sand is allowed to pour out through a hole behind the rear wheels.

21. Two stationary ice skaters, Denise and Lauren, are facing each other and use the palms of their hands to push each other in opposite directions. Denise, with a mass of 50 kg, moves off in a straight line with a speed of 1.2 m s^{-1}. Lauren moves off in the opposite direction with a speed of 1.5 m s^{-1}.

Figure 11.16

   (a) Calculate Lauren's mass.
   (b) Calculate the magnitude of the impulse that results in Denise's gain in speed.
   (c) Calculate the magnitude of the impulse on Lauren while the girls are pushing each other away.
   (d) What is the total momentum of the system of Denise and Lauren just after they push each other away?
   (e) Would it make any difference to their final velocities if they pushed each other harder? Explain.

22. Gavin and Andrew are keen rollerbladers. Gavin approaches his stationary brother at a speed of 2.0 m s^{-1} and bumps into him. As a result of the collision, Gavin, who has a mass of 60 kg, stops moving and Andrew, who has a mass of 70 kg, moves off in a straight line. The surface on which they are 'blading' is smooth enough that friction can be ignored.

Figure 11.17

   (a) With what speed does Andrew move off?
   (b) Calculate the magnitude of the impulse on Gavin as a result of the bump.
   (c) Calculate the magnitude of Gavin's change in momentum.
   (d) Calculate the magnitude of Andrew's change in momentum.
   (e) How would the motion of each of the brothers after their interaction be different if they pushed each other instead of just bumping?
   (f) If Gavin held onto Andrew so that they moved off together what would be their final velocity?

23. An unfortunate driver of mass 50 kg, travelling on an icy road in her 1200 kg car, collides with a stationary police car with a total mass (including occupants) of 1500 kg. The tangled wreck moves off after the collision with a speed of 7.0 m s^{-1}. The frictional force on both cars can be assumed to be negligible.

(a) At what speed was the unfortunate driver travelling before her car hit the police car?
(b) What was the impulse on the police car due to the collision?
(c) What was the impulse on the unfortunate driver of the offending car (who was wearing a properly fitted seatbelt) due to the impact with the police car?
(d) If the duration of the collision was 0.10 s, what average net force was applied to the police car?

24. A car of mass 1500 kg travelling due west at a speed of 20 m s^{-1} on an icy road collides with a truck of mass 2000 kg travelling at the same speed in the opposite direction. The vehicles lock together after impact.
(a) What is the velocity of the vehicles immediately after the collision?
(b) Which vehicle experiences the greater change in speed?
(c) Which vehicle experiences the greater change in momentum?
(d) Which vehicle experiences the greater force?

25. A train of mass 4.0×10^6 kg rolls freely along a horizontal track at a speed of 3.0 m s^{-1} towards a loaded coal cart. The mass of the coal cart is 5×10^5 kg and it is rolling freely at a speed of 2.0 m s^{-1} in the same direction as the train. When the train reaches the coal cart, they remain in contact and continue to roll freely. What is their common speed after contact is made?

26. In a real collision between two cars on a bitumen road on a dry day, is it reasonable to assume that the total momentum of the two cars is conserved? Explain your answer.

27. A well-meaning politician makes the suggestion that if cars were completely surrounded by rubber 'bumpers' like those on dodgem cars, they would simply bounce off each other in a collision and the passengers would be safer. Discuss the merit of this suggestion in terms of impulse, change in momentum and force.

28. In a paragraph, discuss the accuracy of the following statement. Make estimates of the physical characteristics of the car and the wall so that you can support your arguments with calculations.

> When a car collides with a solid concrete wall firmly embedded into the ground, the total momentum of the system is conserved. Therefore, the concrete wall moves, but not quickly enough to allow any measurement of the movement to be made.

29. Design a spreadsheet to model head-on collisions between two cars on an icy road. Assume that the cars are locked together after impact. Use your spreadsheet to predict the speed of the cars after the collision for a range of masses and initial speeds.

11.1 IMPULSE AND CHANGE IN MOMENTUM

Aim

To compare the change in momentum of an object with the impulse delivered by an external force

Apparatus

low-friction trolley or linear air-track glider
timing and recording device (e.g. ticker-timer, spark generator, photogates, motion detector and computer interface)
pulley
light string
load (500 g or 1.0 kg mass)
metre rule

Theory

The impulse delivered to an object by a force is the product of the force and the time interval during which the force acts on the object. The impulse delivered by the net force on an object is equal to the change in momentum of the object.

Method

1. Connect a load of known mass to a dynamics trolley or linear air-track glider with a light string over a pulley, as shown in figure 11.18.
2. Use your timing and recording device to collect data that will allow you to determine the instantaneous velocity of the trolley or glider at two separate instants as the load is falling.
3. Measure and record the mass of the trolley or glider.

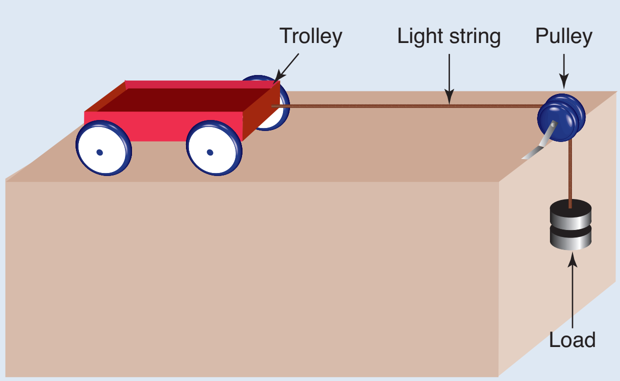


Figure 11.18

Analysis and questions

Use your record of the motion to determine the instantaneous velocity at two separate instants and hence calculate the change in velocity.

1. What is the mass of the system?
2. Calculate the change in momentum of the system.
3. What is the magnitude of the net force applied to the system?
4. Use the net force and the appropriate time interval to calculate the impulse delivered to the system by the net force.
5. Compare the impulse and change in momentum of the system, and discuss any difference between your expected results and your calculations.
6. Express the discrepancy between the change in momentum and the impulse as a percentage of the impulse.
7. Which of the measured quantities was the least accurate? Why?

11.2 SIMULATING A COLLISION

Aim

To show that momentum is conserved in a collision in which there are no unbalanced external forces acting on the system

Apparatus

low-friction trolleys or linear air-track gliders
timing and recording device (e.g. ticker-timer, spark generator, photogates, motion detector and computer interface)
brick or other weight to add to one trolley or glider
balance suitable for measuring the mass of the trolleys or gliders and the added weight
velcro, double-sided tape or plasticine
metre rule

Theory

If the net force acting on a system is zero, the total momentum of the system does not change. This statement is an expression of the Law of Conservation of Momentum. Therefore, if no external forces act on two vehicles during a collision between them, the total momentum of the system of the two vehicles remains constant. It follows that the change in momentum of the first car is equal and opposite to the change in momentum of the second car.

Method

1. Use two low-friction trolleys or linear air-track gliders to simulate a collision between a furniture truck and a medium-sized passenger car. The truck, travelling at a moderate speed, collides with the rear end of the car on an icy road. After the collision, the two vehicles lock together.
2. Attach an adhesive substance (e.g. velcro, double-sided sticky tape or plasticine) to one or both trolleys or gliders to ensure that they lock together after the collision. Record the mass of each 'vehicle' (after adhesive is attached) before setting up the collision. Place a small, light, unrestrained object on to each of the vehicles to represent loose objects.
3. Use your timing and recording device to collect data that will allow you to determine the velocity of each 'vehicle' just before and just after the collision.

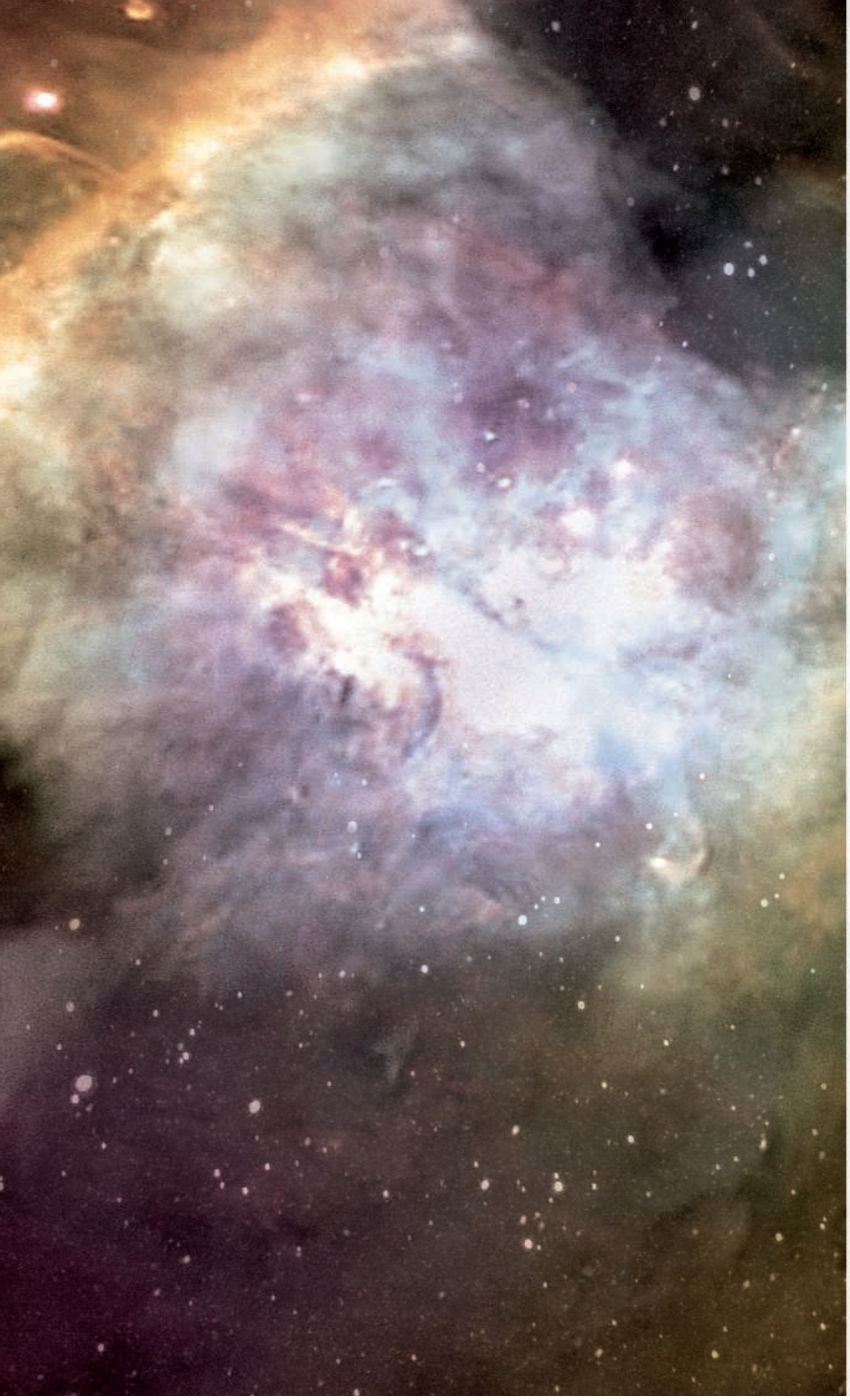
Results

Record your data in a table similar to the one below.

Analysis and questions

1. Describe the motion of each of the loose objects after the collision. What are the implications of your observations for the drivers and passengers in each vehicle?
2. What was the total momentum of the system before the collision?
3. If there were no unbalanced external forces acting on this system, what would you expect the total momentum to be after the collision?
4. What was the total momentum of the system after the collision?
5. How do you account for the fact that momentum was not fully conserved in this collision? Mass was recorded in the tables in grams. Why is there no need to convert it to kilograms?
6. What was the impulse applied to the car during the collision?
7. What was the impulse applied to the furniture truck during the collision?
8. According to Newton's Third Law of Motion, the impulse applied to the car by the furniture truck should be equal to the impulse applied to the furniture truck by the car. Explain your answers to questions 6 and 7 in the light of this.
9. An elastic collision is one in which the total kinetic energy of the system before the collision is the same as the total kinetic energy after the collision. Is this simulated collision elastic? Show your reasoning.

	FURNITURE TRUCK	MEDIUM-SIZED CAR
Mass (g)		
Velocity before collision (cm s^{-1})		
Velocity after collision (cm s^{-1})		
Momentum before collision (g cm s^{-1})		
Momentum after collision (g cm s^{-1})		



PRELIMINARY MODULE

THE COSMIC ENGINE

CHAPTER 12 THE BIG-BANG COSMOLOGY

Figure 12.1 *The Hubble deep field, a star field revealing many galaxies. How many can you pick out?*

Remember

Before beginning this chapter, you should be able to:

- understand that energy may be released from the nuclei of atoms
- describe the features and locations of protons, neutrons and electrons in the atom
- describe some of the difficulties of obtaining information about the universe.

Key content

At the end of this chapter you should be able to:

- outline the historical development of models of the universe, and assess one of those models
- outline the prediction of the expansion of the universe by Friedmann and its subsequent discovery by Hubble
- describe the probable origins of the universe
- describe the Big Bang Theory, including the transformation of radiation into matter that followed
- identify that the equivalence of energy and mass was first described by Einstein
- outline how the accretion of galaxies and stars occurred.

How big is the universe? How old is the universe? How did the universe begin? Seeking the answers to these questions is the life-long task of today's cosmologists. It is a humbling task because each small question answered along the way reveals even more unknowns and more questions to be answered. Yet, these same cosmologists are very aware of how far our view of the universe has come in the past two thousand years, and especially in the past one hundred years.

This chapter looks at the development of models of the universe, beginning with Aristotle who lived more than two thousand years ago. We then focus on the developments of the twentieth century before reviewing the current theory of the beginnings of our universe, known as 'the big-bang cosmology'.

12.1 OUR VIEW OF THE UNIVERSE

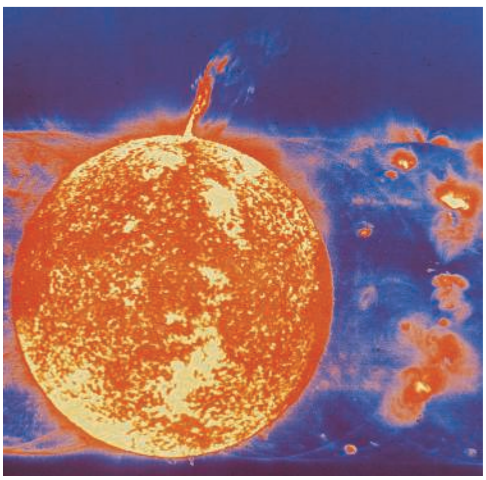


Figure 12.2 *The Sun, revealing the granular nature of its surface as well as flares and spicules*

Our Sun is a star, the closest star to us. It is a huge ball of gas (about 80% hydrogen and 20% helium) that produces its energy by the nuclear fusion (or joining) of the hydrogen nuclei within its core. This is the very same nuclear reaction that takes place inside the detonation of a thermonuclear 'hydrogen' bomb. In fact, the energy emitted from the Sun in just one second would run a 1 kW electric heater for over one hundred thousand million million years!

As remarkable as all this might seem, our Sun is just an average, middle-aged, yellow star, not unlike 100 billion other stars located within our galaxy, the Milky Way.

A galaxy is a large collection of stars and when you look at the night sky and watch the stars, you are looking at our own galaxy. Every individual star you can observe is a member of our galaxy. When you look at the bright strip across the sky called the Milky Way, you are looking toward the centre of our galaxy. The shape of our galaxy has been deduced to be a rotating 'barred spiral' galaxy, as shown in figure 12.3.

Figure 12.3 *A barred spiral galaxy not unlike our own*

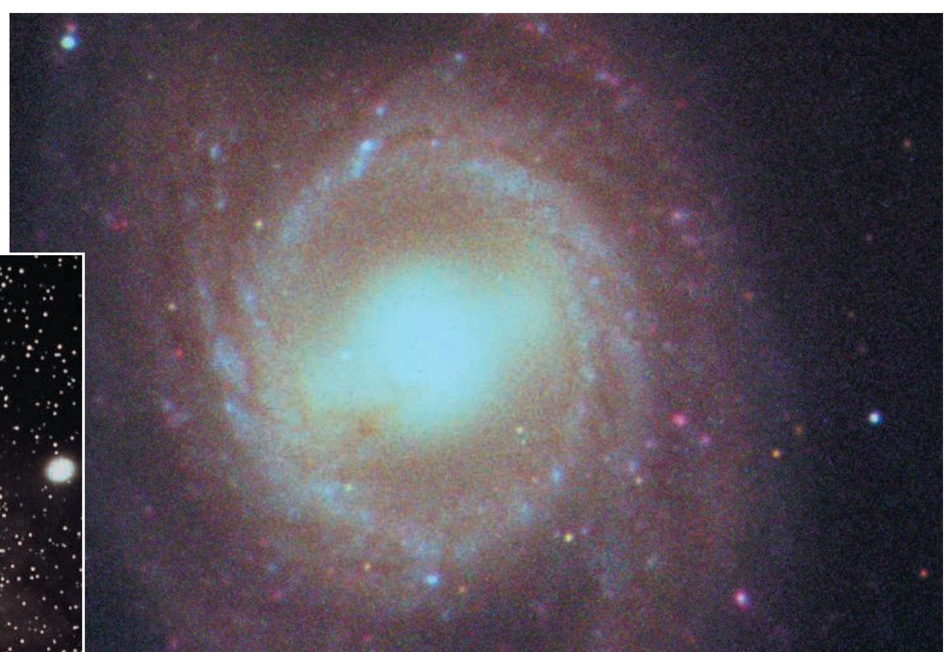
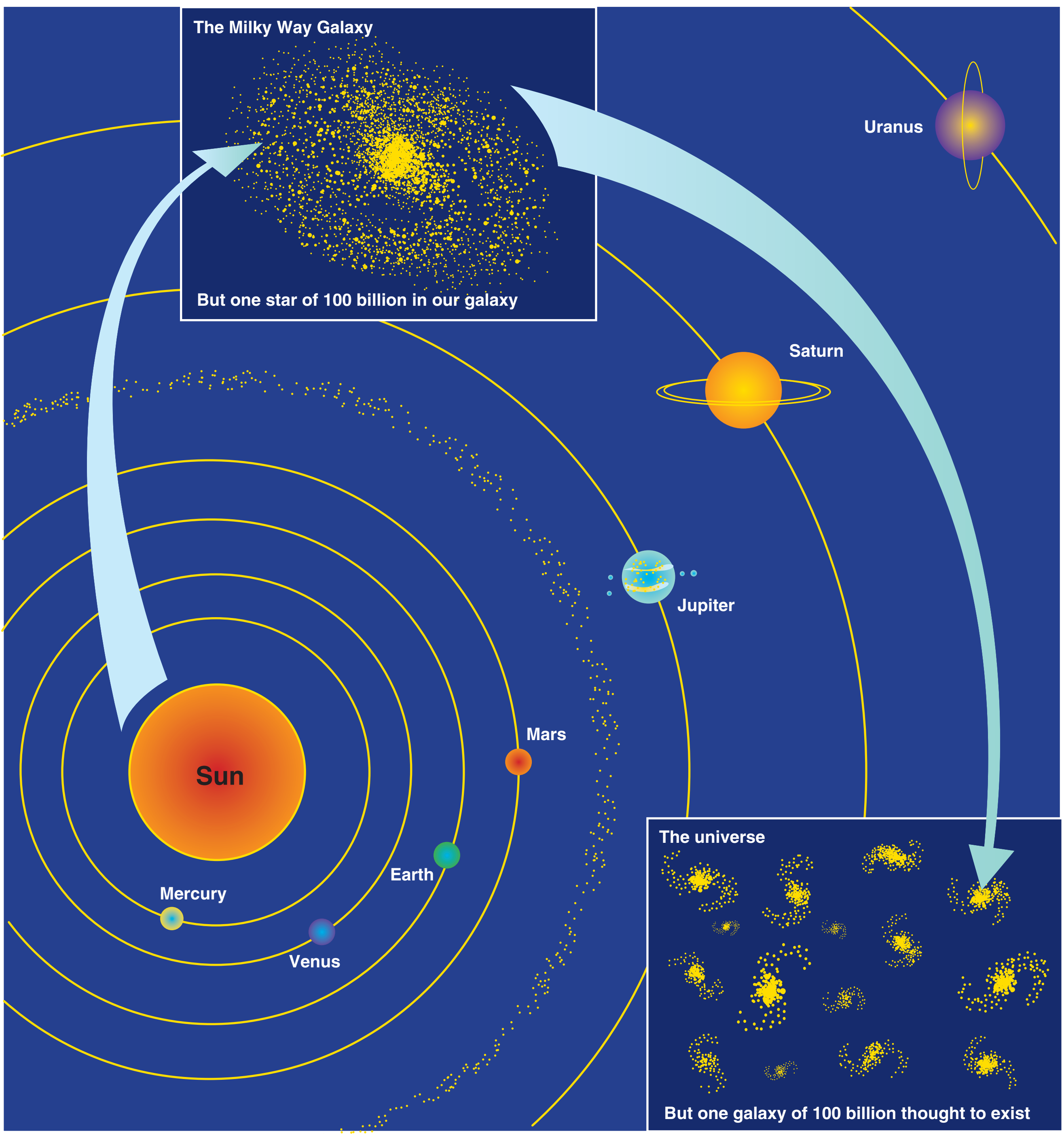


Figure 12.4 *A neighbouring galaxy, Andromeda*

Our galaxy is not alone in space. Figure 12.1 on page 250 shows a very small patch of sky that is rich in galaxies, photographed with the Hubble telescope. There are many billions of galaxies that can be observed around the sky.

In all it is a very humbling picture: we live on a small blue planet, just one of a set of eight that orbit an average star. That star is just one of a set of 100 billion other stars that make up our galaxy, which is in turn just one of an estimated 100 billion galaxies in the universe. It is a view that underlines our insignificance in the vastness of space and is illustrated in figure 12.5.

Figure 12.5 *The vastness of space . . .*



12.2 HISTORICAL DEVELOPMENT OF MODELS OF THE UNIVERSE

The ancient Greeks

More than 2300 years ago lived one of the most influential thinkers the world has ever known. The Greek philosopher Aristotle, who lived from 384 to 322 BC, used a logical cause-and-effect approach to explain physical events. For example, against prevailing opinion he reasoned that the Earth was round rather than flat and provided three different proofs to back up his claim.

Figure 12.6 Aristotle as portrayed by Rembrandt

However, he also believed that the Earth was at the centre of the universe and that the Sun, the Moon and the visible planets, as well as a celestial sphere containing all the stars, revolved around the Earth. This belief may seem strange to us now; however, Aristotle had no technology with which to examine the heavens, other than his eyes. To the unaided eye it certainly looks as though the stars, the Sun and planets are revolving around the Earth, and it certainly does not feel as though the Earth is moving at all. This type of model, with the Earth placed at the centre, is known as a **geocentric model**. Aristotle's model used a system of 55 transparent concentric spheres rotating around the Earth to explain the observed motion of the stars and planets. It is a measure of the esteem in which Aristotle was held that it took science two thousand years to cast aside this model of the universe.

A geocentric model of the universe has the Earth placed at its centre.

Aristotle also believed that the Earth and all things upon it were made up of four elements: fire, air, earth and water. The heavens were believed to have been made of a fifth element called quintessence.

In about 240 BC, another Greek, Aristarchus, put forward an alternative view. He suggested that:

- the Sun is much bigger than the Earth
- the Sun is at the centre of the universe and the Earth orbits it
- the Earth rotates on an axis once per day, producing the apparent motion of the Sun and stars.

A model such as this, with the Sun at the centre, is known as a **heliocentric model**. Aristarchus' model, remarkably close to the truth as we know it, did not gain favour because it was not sufficiently detailed to allow predictions in the manner of Aristotle's model.

A heliocentric model of the universe has the Sun placed at its centre.

Figure 12.7 *Ptolemy and his geocentric model of the universe*

In AD 140, Aristotle's geocentric model of the universe was developed and refined by Ptolemy, the last of the ancient Greek astronomers. He contrived an elaborate model of circles within circles (see figure below) that was remarkably successful in predicting the observed motions of the heavens — so successful that it was later adopted by the Roman Church as the 'correct' picture of the universe. With the weight of the church behind it, this model predominated over all others for about 1400 years!

Nicholas Copernicus (1473–1542)

The situation began to change in the early 1500s. In 1514, a Polish churchman called Copernicus proposed that the Sun was stationary at the centre of the universe and that everything else revolved in circles about it (see figure below). In his readings he had certainly encountered the earlier suggestion of Aristarchus, that the Earth revolved around the Sun.

Copernicus' discussions on the matter were done quietly, so as not to raise the ire of the church, but in 1543, shortly before his death, he formally published his ideas. As expected, his work was branded as heretical; that is, a crime against the teachings of the Church. This action slowed the acceptance of Copernicus' model because, in the time of the inquisition, this had become a risky theory for scholars. In any event, with the available technology of the time, Copernicus' model was no more successful in predicting the motion of the heavens than Ptolemy's model.

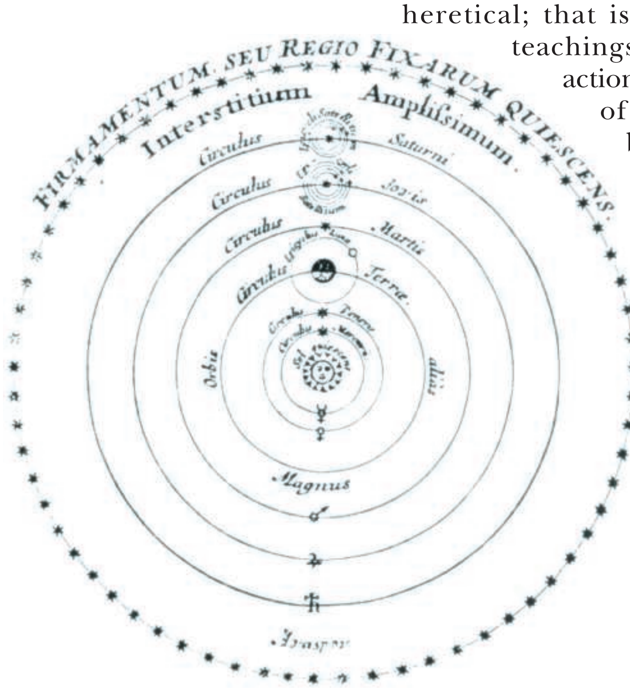


Figure 12.8 *Copernicus and his heliocentric model*

Tycho Brahe (1546–1601)

During the late 1500s, Danish nobleman and astronomer Tycho Brahe studied and plotted the night sky with meticulous care and accuracy. This was an extraordinary achievement because he had no telescope to help him — telescopes hadn't been invented yet! Instead, he designed and constructed his own astronomical measuring instruments, as shown in figure 12.9, and instigated a regimen of regular and methodical nightly observations. His methods were so thorough that he was able to make measurements as accurate as 0.5 arc minutes (30 arc seconds), while other astronomers were limited to approximately 15 arc minutes. His comprehensive knowledge of the stars allowed him to observe and record two supernovae during his life, quite an accomplishment even for modern astronomers.

Brahe's own model of the universe was a combination of the geocentric and the heliocentric models. His model had all of the planets (except Earth) revolving around the Sun, while the Sun revolved around a stationary Earth. Brahe devised this model because he found it impossible to accept that the Earth moved. Again, the technology of the time failed to show any evidence of a moving Earth. The evidence sought would have been a slight shift in the positions of some stars as the Earth orbited the Sun. This effect is called parallax, and the largest parallax of any star is less than one second of arc, or less than a thirtieth of Brahe's best measurements. Brahe had constructed the most sophisticated observatory of his day, and yet he could detect no parallax effect. Therefore, it is not surprising that he found it difficult to totally reject the geocentric model.

Figure 12.9 *(a) Tycho Brahe and one of his instruments and (b) Tycho Brahe*

eBook*plus*

Weblink:
Parallax

Despite this, Brahe's supernovae observations were enough to shake the Ptolemaic view that the stars were unchanging, and his real legacy to the world of astronomy was his years of recorded observations.

PHYSICS FACT

Tycho Brahe was an unusual scientist — a nobleman of some means, who possessed a silver nose (just visible in figure 12.9b). His own nose had been separated from his face in a duel with another nobleman. It is said he had to carry around a pot of glue in case the metal nose fell off!

Johannes Kepler (1571–1630)

In the last few years of his life, Tycho Brahe worked in Prague. His assistant was a clever mathematician called Johannes Kepler, who had already published a theory on planetary orbits based on geometric shapes. When Brahe died, it was Kepler who inherited his wealth of meticulously gathered data. For Kepler, a believer in the Copernican model, the raw data was just what he needed to find a mathematical basis for the motion of the planets. He worked hard for many years, applying Brahe's data to Copernicus' idea and eventually, in the early 1600s, produced an

12.1
Ellipses

improved heliocentric model of the universe. His model said that the planets moved about the Sun, not in circles but in ellipses, and the mathematics of it was encapsulated in three laws:

1. The Law of Ellipses: each planet moves in an ellipse with the Sun at one focus. In practical activity 12.1 you will learn more about the geometry of ellipses.
2. The Law of Areas: the speed of the planets along their elliptical orbits is such that they sweep out equal areas in equal periods of time. In effect, this means that the closer the planets are to the Sun the faster they travel along their orbit.
3. The Law of Periods: the period, T, of the orbit of a planet (that is, the time it takes to complete an orbit) is related to the average radius, R_{av}, of the orbit as follows:

$$\frac{T^2}{R_{av}^3} = \mathrm{k}$$

where

k = a constant.

Note that if the period, T, is in years and the radius, R, is in AU (1 astronomical unit = the average radius of the Earth's orbit), then k = 1. Furthermore, when comparing any two planets A and B, it can be said that:

$$\left(\frac{T^2}{R_{av}^3}\right)_{\text{planet A}} = \left(\frac{T^2}{R_{av}^3}\right)_{\text{planet B}}$$

This law allowed Kepler to make calculations of the relative distances of the planets in the solar system. Some of his work is shown in figure 12.10.

With this new model, Kepler was able to predict the motion of the planets with greater accuracy than Ptolemy's 1500-year-old geocentric model. For the first time, there were good, scientific reasons for scholars to adopt the heliocentric model. Being a protestant in Austria who also enjoyed some favour with the royal court, Kepler was able to indulge his belief in the Copernican model without placing himself at great risk. However, the same situation did not exist closer to Rome.

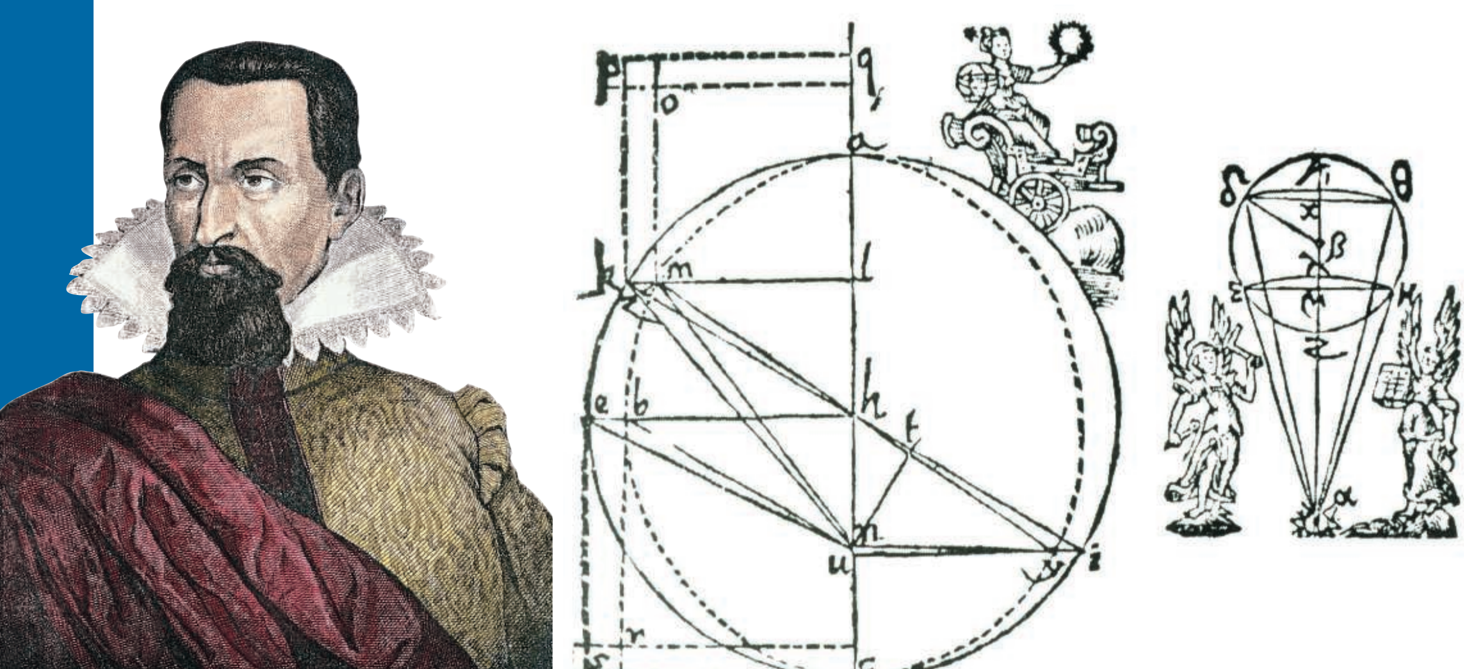


Figure 12.10 *Johannes Kepler and his elliptical heliocentric model*

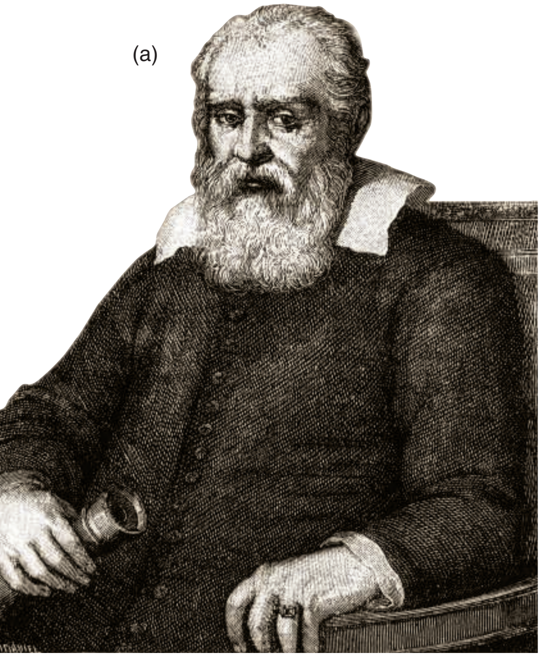


Figure 12.11 ***(a) Galileo Galilei and one of his telescopes and (b) Galileo's sketches of the Moon***

Galileo Galilei (1564–1642)

In 1600, Giordano Bruno, a well-known Italian philosopher and physicist, was burned at the stake by the Holy Inquisition. His crime was that he believed that the Earth moved around the Sun and that there might be inhabited worlds in the universe other than our own.

These were difficult times for Italian physicists who believed in the Copernican model, and yet this was precisely the situation in which Galileo Galilei found himself. Once professor at the university of Pisa, Galileo had a profound influence over many areas of physics, and today is revered as one of the most important people in the history of physics. His work in astronomy was no less profound.

In 1608, the telescope had been invented in Holland (although there is some controversy over who the inventor actually was). By the following year, Galileo had heard about telescopes and began making his own. He became the first person to point one at the night sky. What he saw must have staggered him. If you have never done so, try taking a pair of binoculars and go out and look at the night sky. What strikes everyone when they do this for the first time is the sheer volume of stars that can be seen, stars that were not visible before because they are too faint to be seen with the naked eye.

Galileo made many discoveries with his telescope, not the least of which was that the planet Jupiter had four observable moons that orbited it. He even tracked and plotted their rotations about the planet. The significant point, however, was that the moons were orbiting Jupiter, not the Earth. This was proof that Ptolemy's complicated geocentric model was incorrect. It was, in fact, evidence in favour of the Copernican (heliocentric) model and that was bound to cause trouble for Galileo.

In 1616, the Church issued a warning to Galileo to desist from his public advocacy of the Copernican model. He did not; instead continuing to publish his views. In 1632, he had published a book for the layman, written in Italian rather than the more scholarly Latin, called *Dialogue Concerning the Two Chief World Systems*. Finally, in 1633, he was tried by the Inquisition and found guilty of holding and teaching the Copernican doctrine. He was placed under house arrest for the remaining nine years of his life and allowed very few visitors, so dangerous were his teachings believed to be. But the works of both Kepler and Galileo had dealt a deathblow to the Earth-centred model, and there was no going back.

Isaac Newton (1643–1727)

There is no doubting that Sir Isaac Newton was another of the most important people of physics, a genius whose influence has been felt in many varied fields within physics. In 1687, he published an enormously important work titled *Philosophiae Naturalis Principia Mathematica*. It contained theories on the motions of objects as well as a new type of mathematics (calculus, though he called it 'fluxions') needed to analyse these motions.

To many, Isaac Newton is the fellow who had an apple fall on his head and then discovered gravity. It is doubtful that this incident ever actually occurred; however, the story does serve to highlight another of Newton's achievements. Of course, the idea that objects fall was known before Newton; however, he realised that the force of gravity that acts on (say) an apple close to the Earth's surface, was the same force that held the Moon in its orbit about the Earth, and the Earth in its orbit around the Sun.

In 1684, Edmond Halley (of 'Halley's comet' fame) proposed that the force that acted between the Sun and the planets, whatever its nature (he didn't know what the force was), was inversely proportional to the square of the distance of the planet from the Sun. This means that a planet twice as distant as another would experience just one-quarter of the attractive force. Expressed as an equation:

$$F \propto \frac{1}{d^2}.$$

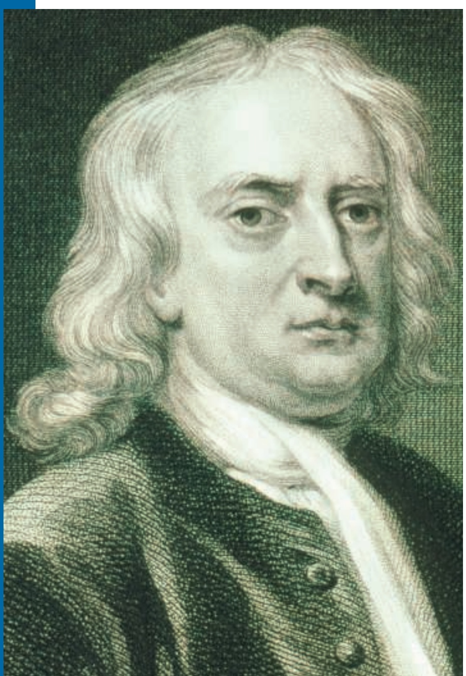


Figure 12.12 *Sir Isaac Newton*

Newton used this idea, along with Kepler's laws and his own idea about gravity, to deduce his Law of Universal Gravitation, which describes the force of gravity that exists between any two masses, but especially large ones such as planets. This law states that the gravitational force of attraction, F, between two masses, m_1 and m_2, is proportional to the product of the masses and inversely proportional to the square of their separation distance, r. That is:

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

where

G = a constant.

The Law of Universal Gravitation placed the physics of planetary orbits on a firm mathematical footing. From this one law, Halley's inverse square relationship is evident and Kepler's laws can be derived, providing an explanation as to why the planetary orbits are ellipses and not circles (Kepler himself never knew why). In addition, the strength of the gravitational field at the surface of a planet can be deduced. (This is discussed in the 'Space' module of the Physics HSC course.)

The physics of Isaac Newton is tremendously successful in explaining the way the physical world works and it dominated astronomy right up to the twentieth century.

12.3 AN EXPANDING UNIVERSE

The mechanistic heliocentric model of the universe established by Copernicus, Galileo, Kepler and Newton did not change during the eighteenth and nineteenth centuries. However, early in the twentieth century, a new breed of physicists equipped with fast-improving technology began to modify and extend the model.

Albert Einstein (1879–1955)

In 1905, a young physicist named Albert Einstein, working almost entirely on his own, published the first ideas that were to become his Special Theory of Relativity. This was a new way to look at the fabric of the universe: a space–time continuum in which time can slow down and objects shrink if they are going fast enough, although nothing can exceed the speed of light.

The theory predicted that if you try to accelerate an object up to the speed of light, then some of the energy you put into the object is actually converted into more mass for the object, making it bigger and harder to accelerate, so that you can never have enough energy to reach the speed of light. Therefore, Einstein suggested that energy can be converted to mass, and vice versa, according to the equation $E = mc^2$. This *energy/mass equivalence* is one of the most important outcomes of the theory of relativity.

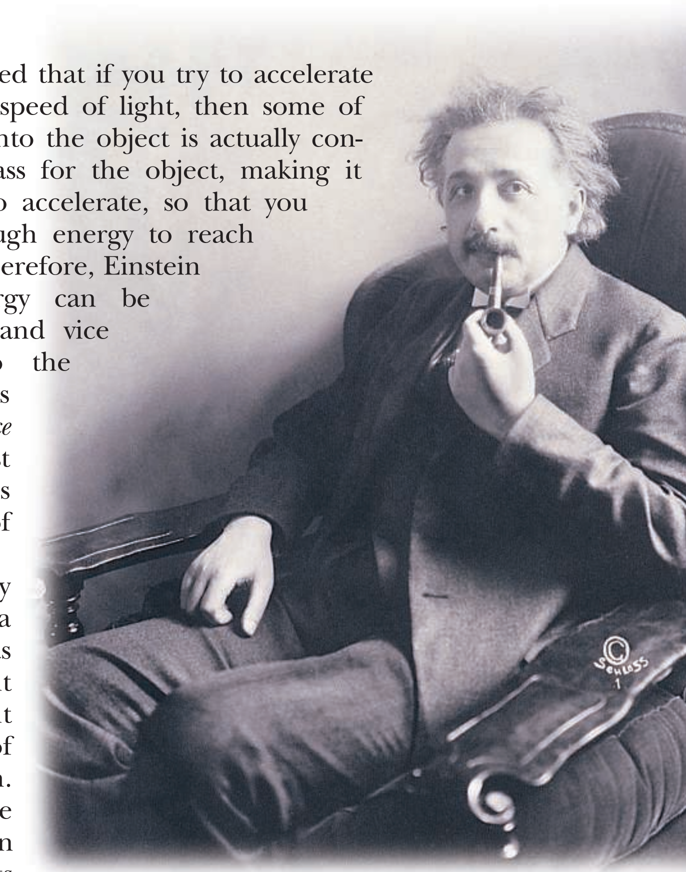


Figure 12.13 *Albert Einstein in 1920*

eBook*plus*

Weblink:
Einstein image and impact

The Special Theory of Relativity was a success but there was still a problem — it was not consistent with Newton's Law of Universal Gravitation. This stated that the force of gravity between two celestial objects (for example, the Sun and the Earth) depended on the mass of the objects and the distance between them, and that is all. This meant that gravity would act instantaneously. If the Sun were to suddenly disappear then so too would the force holding the Earth — instantly! However, the Earth would still have eight minutes left before the Sun's light disappeared (because it lies eight light-minutes from the Sun).

Einstein said that nothing could happen instantaneously like that. He said it could happen no faster than the speed of light, and so for the next ten years he worked at including gravity into his theory. The result, published in 1915, was the General Theory of Relativity.

This theory presented a new way of thinking about gravity, denying that it is a regular force at all. Rather, it is simply an effect, the result of warped space–time. Large masses have the ability to warp space–time, and space–time affects the way the masses move. Planets that appear to be moving in a curved path in space (that is, orbiting around stars) are actually following straight paths through curved space–time.

There were many unexpected outcomes of this new theory, outcomes that were not testable using the technology of the time. However, as the twentieth century progressed, technology improved quickly, and eventually it became possible to design and conduct experiments that could test the theory. Each experiment conducted verified the predictions. The following is a short list of examples:

- The orbits of the planets were predicted to precess; that is, the long axis of the elliptical orbit would itself slowly rotate about the Sun. Mercury has the most eccentric orbit of the planets and so it should precess the most. This had already been noticed prior to 1915, so it served immediately as evidence. Since then, the precession of the other planets has been verified as well.

- The theory predicts that even light rays should be affected by warped space–time and should curl around stars. This effect, called gravitational lensing, was verified in 1919 by carefully plotting the observed positions of stars from West Africa during a total eclipse. Recently, gravitational lensing of distant quasars has been studied by astronomers using the Hubble Space Telescope to try to measure the size of our universe.
- The General Theory of Relativity predicts that time should run slower near an object as large as a planet. This prediction was tested and verified in 1962.

Einstein had presented a revolutionary way to look at the universe, especially considering the time it was published — in the middle of World War I. It is regarded by physicists generally as one of the two great theories of the twentieth century (the other being quantum mechanics). However, there was one point on which Einstein disagreed with his own theory. The theory predicts that the universe is expanding, but Einstein believed very strongly that the universe was static and unchanging. He introduced a special 'cosmological constant' into his equations to ensure this result. But here Einstein himself was to be proven wrong, though his General Theory of Relativity was verified experimentally yet again.

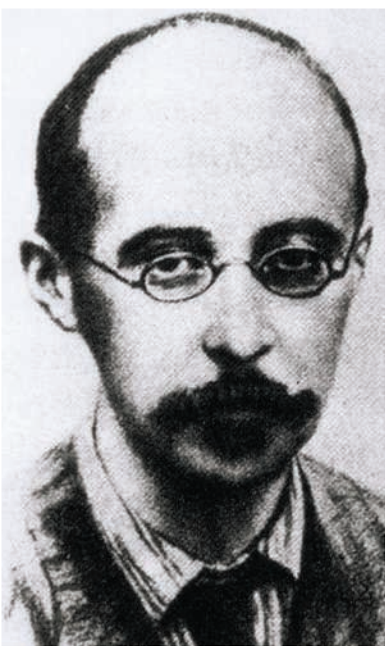


Figure 12.14 *Aleksandr Friedmann*

Aleksandr Friedmann (1888–1925)

Russian Aleksandr Friedmann was professor of mathematics at the University of Leningrad when he read of Einstein's problem with relativity's expanding universe, and how Einstein had resolved the problem. Friedmann considered this an error and he set about proving that the universe must be expanding, based on the theory of general relativity. To do this he had to make two assumptions:

1. that the universe appears identical whichever direction you look
2. this is true whatever your viewpoint in the universe.

Friedmann expressed this expansion of the universe in mathematical terms, saying that the radius of curvature of space increases with time. He published his work in a 1922 article titled 'On the curvature of space' in the German physics journal *Zeitschrift für Physik*. Some correspondence occurred between Einstein and Friedmann before Einstein was prepared to accept that Friedmann was correct. (Friedmann's findings were independently verified by discoveries made by Belgian astronomer Georges Lemaitre in 1927. A few years later Lemaitre went further and suggested that the universe started out as a single 'space particle'.)

There are, in fact, three different types of expanding universe models that satisfy Friedmann's assumptions, although he only found one of them.

1. In the type that Friedmann discovered, called a closed universe, the mutual gravitation of all of the matter in the universe is able to bring the expansion to a halt and then pull all of that matter back together, causing the universe to contract again.
2. In the second type of model, called a flat universe, the expansion of the universe is just fast enough to balance out the force of gravity, so that the expansion slows but never quite stops. This model was first developed by Dutch astronomer Willem de Sitter in 1917 from Einstein's equations, so it is known as the 'Einstein–de Sitter universe'.
3. In the third type of model, called an open universe, the universe is expanding so fast that gravity can never stop it.

The most successful modification of the Big Bang Theory, known as Inflation Theory, says that the universe underwent an early period of very rapid expansion before slowing down. This effectively separates the issues of the geometry of space and its expansion. Experimental results from the WMAP space probe survey confirm that our universe is geometrically flat while also expanding at an ever-increasing rate.

Edwin Hubble (1889–1953)

In 1919, Edwin Hubble, a lawyer with a PhD in astronomy, began to work at the Mount Wilson Observatory in California, which then possessed the world's most powerful telescope. He was studying nebulae, thought to be clouds of glowing gas that appear, through small telescopes, as small fuzzy patches in the night sky. With the Mount Wilson telescope, however, Hubble was able to see images of individual stars in these nebulae. In 1924, he noticed a particular type of pulsating star in the Andromeda nebula. This type of star was called a 'Cepheid variable'. Discovered by Henrietta Leavitt in about 1910, these stars are so predictable that their distance can easily be calculated. Hubble performed the calculation and found that Andromeda is 800 000 light years away. This was quite a surprise because at this distance Andromeda was far beyond the most distant stars known. Rather than being a gaseous object within our galaxy, Hubble had shown that Andromeda was really a separate galaxy (some nebulae are clouds of gas; however, prior to Hubble astronomers could not distinguish those that were really galaxies). This was the first evidence to suggest that there are galaxies other than our own, and Hubble went on to discover and classify about twenty-four different galaxies. A picture began to emerge of our galaxy, the Milky Way, as but one of many galaxies that make up the universe.

Hubble's discovery was a surprise to the whole world, not just the world of science. However, he had another surprise to come. In his study of the newly discovered galaxies, he measured the red shift of their light. This red-shift effect had been discovered in 1912 by Vesto Slipher, who had been using the Lowell observatory in Arizona to observe and analyse the light of spiral nebulae. (Remember, he didn't know that they were galaxies.) Using a glass prism to spread the light out into a spectrum, it becomes possible to identify particular lines (specific wavelengths) in the spectrum. What Slipher noticed was that the lines were closer to the red end of the spectrum than expected — hence the term 'red-shifted'. He deduced that this meant that they were moving away from the Earth, and by measuring the degree of red-shifting it was possible to calculate how fast they were moving. Red-shifting is illustrated in figure 12.15.

eBook*plus*

eLesson:
Expansion of the universe
eles-0038

Interactivity:
Expansion of the universe
int-0057

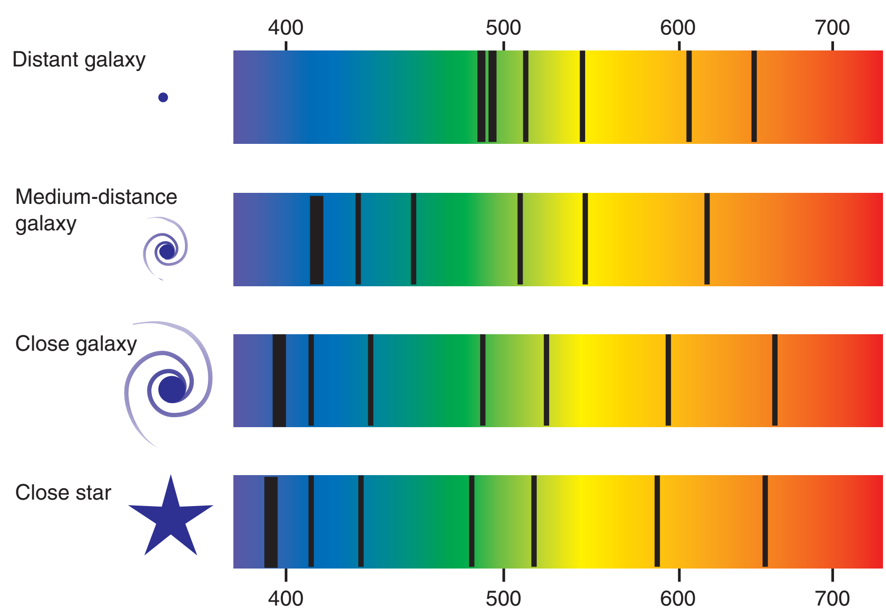


Figure 12.15 *Spectra showing red-shifting*

12.2

The Hubble constant

Hubble applied this same technique to his galaxies. What he found was a surprise to nearly everyone when he published his results in 1929. Almost all the galaxies were red-shifted, meaning that they were moving away from us. Not only that but he discovered a direct relationship between the distance of a galaxy and its apparent speed: the further away a galaxy is, the faster it is receding from us. This has become known as Hubble's Law:

$$v = H_0 D$$

where

v = velocity of recession (km s^{-1})

D = distance from us (megaparsecs, Mpc)

H_0 = Hubble's constant ≈ 70 km s^{-1} Mpc^{-1}.

Note that a parsec is a distance of 3.26 light years, so that one megaparsec (Mpc) equals a distance of 3.26 million light years, or 31 billion billion kilometres. The units 'parsec' and 'megaparsec' are often used to express the enormous distances of space.

Hubble's constant is a measure of the rate of expansion of the universe. It is related to the age of the universe.

The value of **Hubble's constant** was first determined by Hubble to be 500 km s^{-1} Mpc^{-1}. However, the accepted value of H_0 has reduced over the years as improved technology has provided more accurate measurements of the distance to stars and galaxies. In 1999, the Hubble Space Telescope Key Project team — led by Wendy Freedman, Jeremy Mould and Robert Kennicutt — measured the value of Hubble's constant to be 70 kilometres per second per megaparsec, with an uncertainty of 10 per cent.

What Hubble had discovered is that the universe is expanding, just as Friedmann had predicted five years earlier. Moreover, Hubble's expansion was very much like an exploding bomb, with the parts furthest out travelling the fastest. This implied that at some stage in the past all of the matter in our expanding universe, like an exploding bomb, was concentrated at a single place, called a singularity. From this singularity our universe exploded, an event that has become known as 'the big bang'. Using Hubble's constant, it is possible to calculate that the big bang occurred approximately 12 to 14 billion years ago.

The cosmic radiation background

In the same way that a star emits radiation (its light), a young, hot and dense universe should also have emitted radiation, the exact nature of which depended upon its temperature. You can think of it as the glow from the heat of the early universe. But as the universe expanded, like a hot gas that expands, it cooled down. This means that the radiation should also have cooled, although it should still be around today. This was the 1948 prediction of George Gamow, once a student of Friedmann. However, the radiation itself was not discovered until 1965.

Cosmic background radiation is a pattern of background radiation from space that represents the afterglow of the heat of the early universe. Today it is in the form of microwaves.

At this time the inventor of the microwave radiometer, Robert Dicke, suggested that this '**cosmic background radiation**' should by now appear to us as microwaves from space and that a radiometer should be used to search for them. However, before Dicke could follow through on his own suggestion, someone else found his microwaves with a radiometer, by accident.

In 1965, two engineers, Arno Penzias and Robert Wilson, were working on a radiometer built into a telescope that Bell Laboratories was using to track communications satellites. The radiometer was picking up a

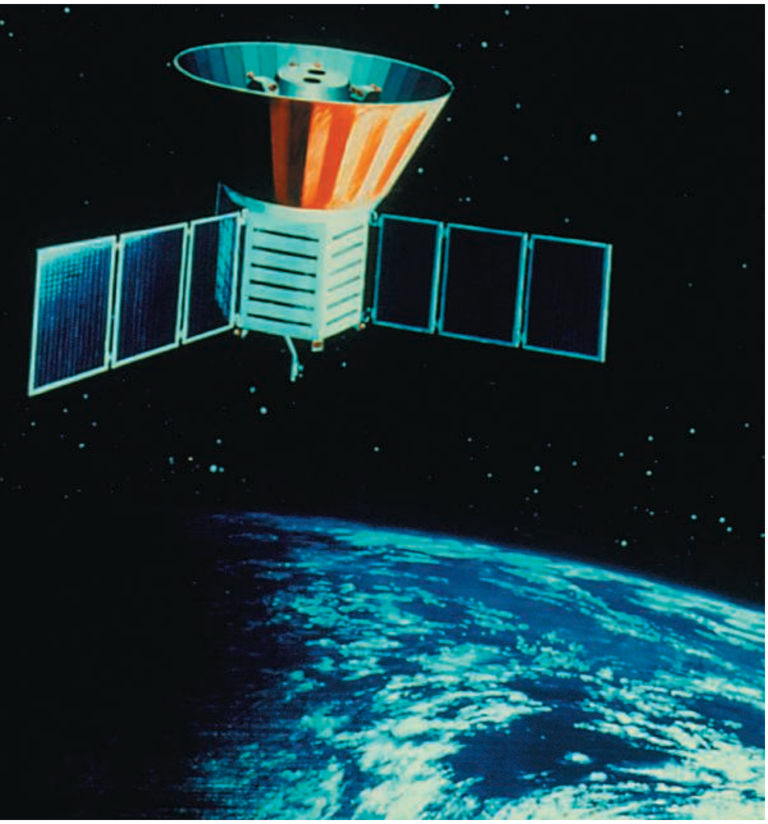


Figure 12.16 ***The COBE satellite. Its measuring instruments are the DIRBE (Diffuse Infra-Red Background Experiment); the FIRAS (Far Infra-Red Absolute Spectrophotometer), which was able to measure the temperature of the background radiation at 2.7 K; and the DMR (Differential Microwave Radiometer), which surveyed the microwave background.***

persistent noise that they were unable to eliminate. It appeared to them that the microwave noise they were detecting came from space — not just from outside the Earth, or the solar system, but outside the galaxy. They believed this because the noise was the same whichever direction the telescope was pointed. When the engineers heard of Dicke's suggestion, they realised that they had inadvertently discovered the cosmic background radiation that Gamow had predicted.

This radiation became the subject of a great deal of research during the late 1980s and early 1990s, and a wide variety of experiments was conducted. Finally, a satellite was launched in 1989 called the COsmic Background Explorer, or COBE. This satellite is shown in figure 12.16. From its orbit around the Earth, COBE surveyed the entire sky, accurately mapping the strength of the radiation.

It was found that the microwave radiation was very nearly the same strength in every direction, varying by just one part in 100 000, which justified Aleksandr Friedmann's original assumption about the uniformity of the universe (see figure 12.17).

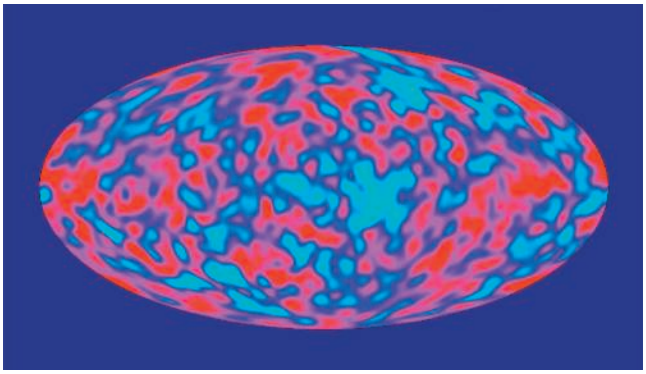


Figure 12.17 ***This computer-enhanced image shows the cosmic microwave background radiation across the entire sky, as surveyed by the COBE satellite. The lumpy appearance is evidence of galaxy formation when the universe was just 300 000 years old.***

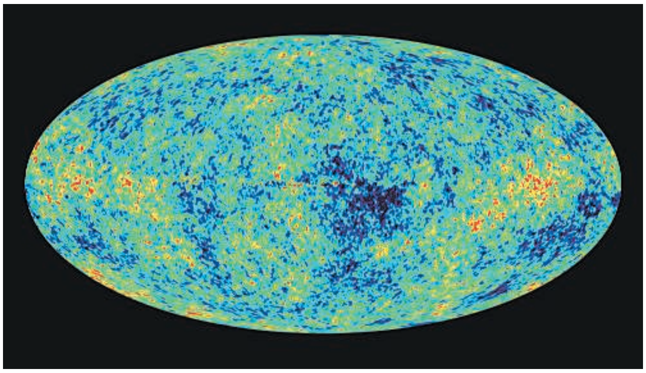


Figure 12.18 ***The computer-enhanced image of cosmic microwave background radiation produced by the WMAP mission. Note that, while similar to the COBE image, its detail is finer, indicating the higher resolution of the instruments on the probe.***

Additionally, COBE included a device that was able to measure the temperature of the background radiation as just 2.7 K, or –273 degrees Celsius! This is precisely what was suggested by Gamow. COBE provided strong evidence that cosmologists were on the right track and more recent evidence has supported this. In 2003, the Wilkinson Microwave

Anisotropy Probe (WMAP) confirmed the COBE results, but to a much higher resolution and accuracy. The study of the data from the probe deduced:

- the age of the universe: 13.7 billion years ± 1%
- the geometry of the universe: flat
- the value of Hubble's constant: $H_0 = 73.5$ km s^{-1} Mpc^{-1} ± 4%
- the content of the universe: 4% atoms, 23% 'cold dark matter', 73% 'dark energy'
- the formation of the first stars: 200 million years after the big bang (earlier than expected).

For more information refer to the WMAP website at http://map.gsfc.nasa.gov.

12.4 THE BIG BANG

As already mentioned, Georges Lemaitre suggested in 1931 that the universe started out as a single 'space particle'. In 1948, George Gamow went further, suggesting that the expanding universe started out as a hot, dense concentration of matter that 'exploded' out to its present form. Fred Hoyle, an eminent English astronomer and science-fiction author, was an opponent of this theory. During a 1950 radio broadcast, he referred to it disparagingly as a '**big bang**'. It was such a catchy title that the name stuck.

The term 'big bang' refers to the rapid expansion of the very early universe.

There was a lengthy debate about the validity of this theory, but evidence slowly gathered in favour of the big-bang cosmology. Then, in 1970, the argument seemed to have been settled. Stephen Hawking and Roger Penrose proved mathematically, based on general relativity, that the universe must have begun with a big bang.

Today, the Big Bang Theory has largely been accepted and is supported by much evidence, such as the cosmic background radiation, the relative proportions of the lighter elements and the Hubble expansion. However, it is not perfect, and modifications continue to be developed to explain unexpected observations. (See Physics in Focus — The dark energy dilemma p. 265.) The current model is one that supports an initial rapid inflation leading to a more slowly expanding, spatially-flat universe that is now starting to accelerate its expansion. The model is known as the Lambda Cold Dark Matter model (or ΛCDM model).

By the application of quantum mechanics to this model, it has been possible to construct a picture of the development of the universe, almost from its start approximately 13.7 billion years ago.

The expansion and cooling

The earliest time that anything at all can be said about the universe, according to quantum physics, was at 10^{-43} seconds after the moment of the big bang itself, which was the beginning of space and time. There simply was no 'before' the big bang because time had not yet begun. There was no 'outside' the big bang because space itself had only just come into being and was much smaller than even the size of a proton, although incredibly hot — about 10^{32} K. All that was present was energy.

eBook*plus*

Weblink:
Big bang theory

At 10^{-35} seconds after the big bang, the universe began a period of rapid exponential expansion, driven by a scalar field of energy filled with massive particles that have been called 'inflatons', which have a repulsive, anti-gravitational effect. Note that inflatons and their field are a proposed mechanism that is yet to be identified.

By about 10^{-34} seconds, the universe had cooled to about 10^{27} K and expanded to about the size of a pea, which represented a tremendous inflation in size. The inflatons had decayed into fundamentals particles, so that much of the original energy had been converted into quarks and electrons, with the remaining energy in the form of photons. (Recall that energy can be converted to mass, according to Einstein's equation $E = mc^2$.)

By 10^{-4} seconds, the quarks had combined to form protons and neutrons. Also present were the electrons, neutrinos and a great deal of radiation that could not move freely without colliding with particles.

By 10^{-2} seconds, the universe had expanded to the size of the solar system, and by the time the universe was one second old it had grown to a size somewhere between one and 10 light years across. At this point it had a temperature of about 10^{10} K and a density of about 10 kg cm^{-3}, and all essential matter particles had been formed (see figure 12.19).

Plasma is an extremely hot and highly ionised state of matter.

Over the next few minutes, the universe, which was in an extremely hot and highly ionised state of matter called **plasma**, began thermonuclear fusion. Within this 'primeval fireball', deuterium, helium and some lithium were formed. (Calculations show that they would have been formed in precisely the proportions observed to exist in the universe today.) The radiation was still trapped within this plasma, so the universe could be said to be opaque.

By the time the universe was 380 000 years old, known as the time of recombination, it had grown to be about one-thousandth of its current size and had cooled to about 3000 K. At this temperature, the electrons could be captured by the hydrogen, deuterium, helium and lithium nuclei to form neutral atoms. Suddenly, the radiation was no longer trapped and it escaped, free to roam the universe. The universe could then be described as transparent and black. This radiation still exists today but has since cooled to 2.7 K, and forms the cosmic microwave background radiation mentioned earlier. Figure 12.19 illustrates the expansion of the universe after the big bang.

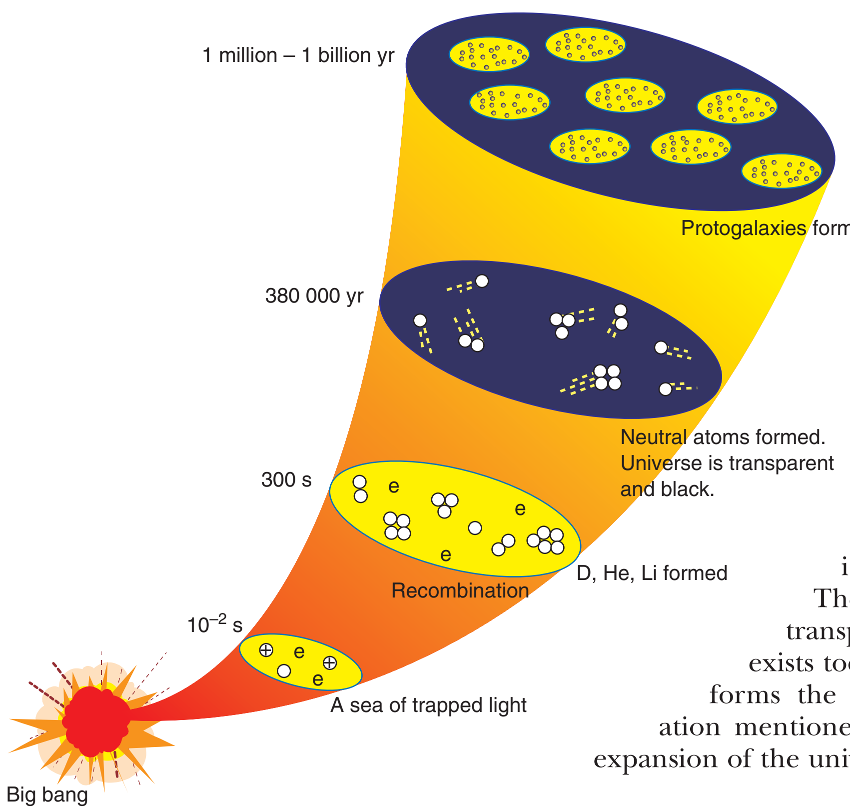


Figure 12.19 *The big bang*

PHYSICS IN FOCUS

The dark energy dilemma

In the mid 1990s cosmologists were confident that they had most of the answers to the beginning and expansion of the universe worked out with just a few mathematical parameters left to measure. The existence of some 'dark matter' that could not be seen had been deduced by watching the rotation of galaxies, but they were confident that the expansion of the universe would be slowing due to the mutual gravitation of the universe's mass.

(continued)

Then, in 1998, a type 1a supernova called 'Albinoni' changed all that. Type 1a supernovae are regarded as 'standard candles' — that is, their actual brightness is known — so by measuring how bright they appear it is possible to calculate how far away they are (this is covered in more detail in chapter 13). The distance to Albinoni worked out to be 18 billion light years! But there were more surprises in store — its red shift was not consistent with this distance. The only way of reconciling this fact was to infer that the expansion of the universe is accelerating.

This finding was entirely contradictory to expectations — there seems to be some unknown anti-gravitational force that is driving the universe apart ever faster. This force appears to be a property of empty space itself, something that we have never before perceived or observed. It was labelled 'dark energy'.

A flurry of activity followed — conferences to brainstorm possible answers and plan experiments to test these ideas. There was general agreement on the Lambda CDM (Cold Dark Matter) model as the simplest universe model that fits the new data. This model mathematically describes a spatially-flat universe made of atoms, cold dark matter and dark energy. Since that time, an international project called ESSENCE set out to observe 200 type 1a supernovae and measure the universe's expansion more thoroughly. The study was concluded at the end of 2007, and the Albinoni result was confirmed.

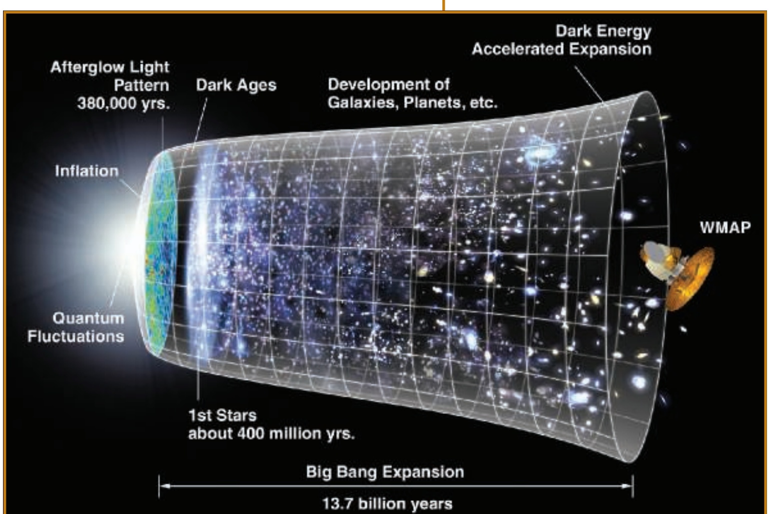


Figure 12.20

In 2001, the WMAP space probe was sent into space to closely survey the cosmic background radiation. The final results released in 2006 yielded some remarkable findings that have supported the Lambda CDM model. See figure 12.20.

From their measurements, the WMAP team have been able to determine the mean density of the universe, which is equivalent to almost 6 protons per cubic metre. The actual density of atoms that we can observe in the universe is about $\frac{1}{4}$ protons per cubic metre. From this information, they have calculated that 96 per cent of the universe is in a form that we have never been able to directly detect, divided up in the proportions shown in figure 12.21.

And so it appears that there is a lot of dark energy around us. It has an effect on astronomical objects, and yet we cannot see it, feel it or otherwise directly measure it. It does have a similar mathematical effect to Einstein's 'cosmological constant' (see page 260), and so it appears to be a property of empty space, with the same value at every point in the universe. However, at the time of writing, cosmologists are no closer to understanding what dark energy is — some suggestions include quintessence, vacuum energy density or perhaps hidden dimensions from quantum string theory. It's shaping up to be cosmology's biggest mystery.

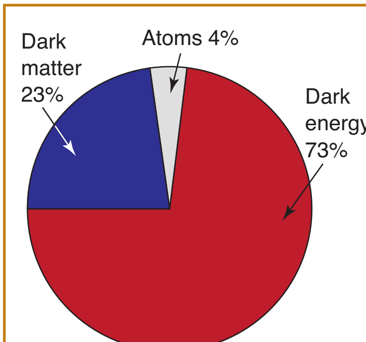


Figure 12.21

12.5 STAR AND GALAXY ACCRETION

The cosmic background radiation images shown in figures 12.17 and 12.18 are valuable to cosmologists because they show two things previously thought to be incompatible in the distribution of matter at the time of recombination. First, they show the uniformity required by the Friedmann expanding universe model. Second, they show a degree of lumpiness. This lumpiness, or uneven texture, was necessary for the formation of stars and galaxies. A huge cloud of perfectly evenly distributed matter can never condense into stars because the pull of gravity is equal in every direction. However, a lumpy cloud has areas of higher density that will experience greater gravitational forces, leading these lumps to grow in density as they gather, or accrete, more matter — a process we call **accretion**.

Accretion **is the process of the growth of a body by gathering or aggregating more matter.**

It is thought that by the time the universe was one-fifth of its present size all of the matter had gathered into discrete gas clouds large enough to be called protogalaxies. It is unclear as yet whether these protogalaxies were the size of galaxy superclusters and later broke up into smaller parts, or whether they were the size of dwarf galaxies and later aggregated into larger groups.

The process of contraction results in an increased rotation. The effect is similar to a spinning ice skater who draws her arms in to her torso in order to spin faster. Most aggregating gas clouds would have had some rotational motion, however small. As the mass within the cloud draws towards the centre, this rotation speeds up. This increased spin also causes the cloud to flatten into a rotating disc shape. This is the most commonly observed shape of galaxies.

The protogalaxies became galaxies as, within them, the process of accretion continued on a smaller scale to form stars. Smaller lumps of gas contracted under mutual gravitation — swirling and collapsing, forming small rotating discs. The centre of such a disc would have continued to become denser and hotter until thermonuclear fusion of hydrogen began within its core so that a new star was formed. The remainder of the disc matter may begin to form into small pieces, which can eventually accrete into planets. The process of star formation was repeated many billions of times within each galaxy, and continues to occur. A type of star known as T Tauri is thought to be undergoing the process just described (see figure 12.22).

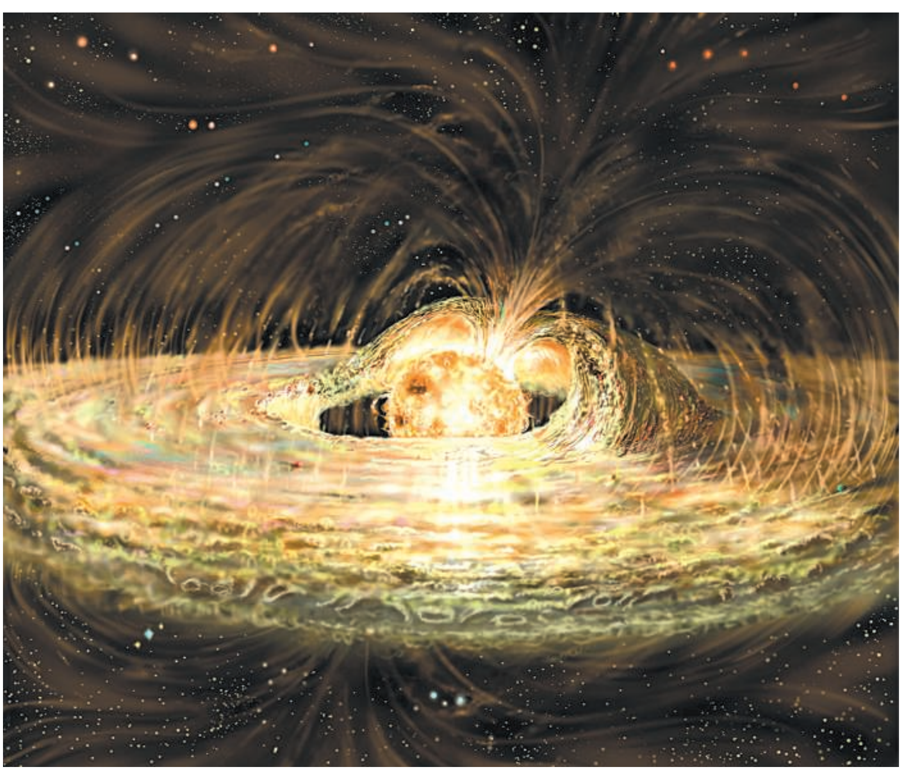


Figure 12.22 *An artist's representation of a T Tauri star, a newborn star forming within a swirling gas cloud. Accreting material falls in along the star's magnetic field lines, crashing in near the pole. The rapid spin of the star also throws material out. This activity causes the light output to vary unpredictably, although the star is normally obscured by the cloud in which it lies.*

SUMMARY

- Aristotle (384–322 BC) established the geocentric (Earth-centred) model of the universe.
- Aristarchus (about 240 BC) suggested the first heliocentric (Sun-centred) model of the universe.
- Ptolemy modified Aristotle's model to improve its agreement with the observed motion in the heavens, in AD 140.
- Nicholas Copernicus (1473–1542) was the next to seriously propose a heliocentric model.
- Tycho Brahe (1546–1601) instigated methodical observation practices to make meticulously detailed recordings of the motion of the planets.
- Johannes Kepler (1571–1630) used Brahe's records to improve Copernicus' model with a mathematical description of the elliptical orbits of the planets. At last, the heliocentric model of the universe was more accurate in predicting the motion of the planets than Ptolemy's geocentric model.
- Galileo Galilei (1564–1642) used a telescope to observe the heavens and noticed detail that contradicted the Aristotlean/Ptolemaic geocentric model.
- Isaac Newton (1642–1727) analysed the force of gravity between two masses to provide a better mathematical description of the workings of the heliocentric model of the universe.
- Albert Einstein (1879–1955) proposed the General Theory of Relativity, which contained much information on space–time.
- Aleksandr Friedmann (1888–1925) demonstrated mathematically that the Theory of Relativity predicted that the universe was expanding.
- Edwin Hubble (1889–1953) discovered that the universe consists of galaxies other than our own and then discovered experimental evidence that the universe is expanding. He described this in a mathematical form known as Hubble's Law.
- The cosmic background radiation, proposed by George Gamow, is the afterglow of a young, hot and dense universe. It was discovered accidentally by Arno Penzias and Robert Wilson after Robert Dicke suggested that the radiation would be in the form of microwaves.
- The term 'big bang' refers to the explosion-like birth of our universe.
- Immediately after the big bang there was only energy, which soon transformed to matter such as quarks and electrons. A great deal of trapped radiation was also present. In the next few minutes, light nuclei were formed by thermonuclear fusion.
- After about 380 000 years, the temperature cooled sufficiently for electrons to bond to nuclei to form neutral atoms. The trapped radiation could then break free. This is called the time of recombination.
- After the time of recombination, galaxies and stars began to form through the process of gravitational collapse and accretion.

QUESTIONS

1. Rule a line across your page. Make a mark at the far left and label it 'Aristotle'. Make another mark at the far right and mark it 'Newton'. Use this to construct a time-line of developing models of the universe.
2. Choose one of the models indicated on the time-line in question 1.
   (a) Describe the model in more detail.
   (b) List the advantages of this model, from the perspective of someone living at that time.
   (c) How did the technology of the time influence or limit the development of this model?
3. List the twentieth-century discoveries that have furthered our model of the universe.
4. Write a one-paragraph description of the current model of the universe. The scale of your description should range from planet to universe.
5. The expansion of the universe was theorised by Friedmann in 1922.
   (a) Even though Einstein acknowledged that Friedmann was correct, state what was still required to verify the theory.
   (b) Identify when this verification occurred, and what form it took.
6. Edwin Hubble's discovery of the expansion of the universe relied upon the earlier work of two scientists, and verified the work of a third. Identify the other scientists and the contribution of each.

7. (a) Describe Hubble's Law in qualitative terms.
   (b) Write down the Hubble's Law equation.
   (c) Explain the significance of the Hubble constant to cosmologists.

8. Explain how Hubble's discovery of the expansion of the universe led to the idea of the big bang.

9. (a) Explain why the discovery of microwave background radiation by Penzias and Wilson was viewed by cosmologists as experimental evidence for the big bang.
   (b) Identify the two scientists upon whose work this interpretation of the microwave background radiation relied.

10. Read the narrative describing the process of the big bang in section 12.4, and then tabulate this information. Your table should include two columns — 'time' and 'brief description'.

11. (a) Describe the significance of the time of recombination.
    (b) Describe why it was not possible for stars or galaxies to form before this time.

12. The processed image from the COBE satellite shown in figure 12.17 is seen as significant by cosmologists because it shows two features previously thought incompatible. Identify these features and describe their significance.

13. Immediately after the big bang there was only energy present, yet the stars are formed of matter.
    (a) Describe the source of this matter.
    (b) State the equation that describes this process.

14. Outline the formation of galaxies through the process of accretion.

15. Compare the formation of stars (and their possible planetary systems) to the formation of galaxies.

12.1 ELLIPSES

Aim

To draw ellipses representing the orbits of the planets of our solar system

Apparatus

a length of string
a drawing board or sheet of thick cardboard
two drawing pins

Theory

An ellipse has the dimensions shown in figure 12.23a. Points F_1 and F_2 are called the focuses of the ellipse. Note that the eccentricity of the ellipse, which is a measure of its elongation, is defined by the ratio $\frac{c}{a}$.

Eccentricity, $e = \frac{c}{a}$

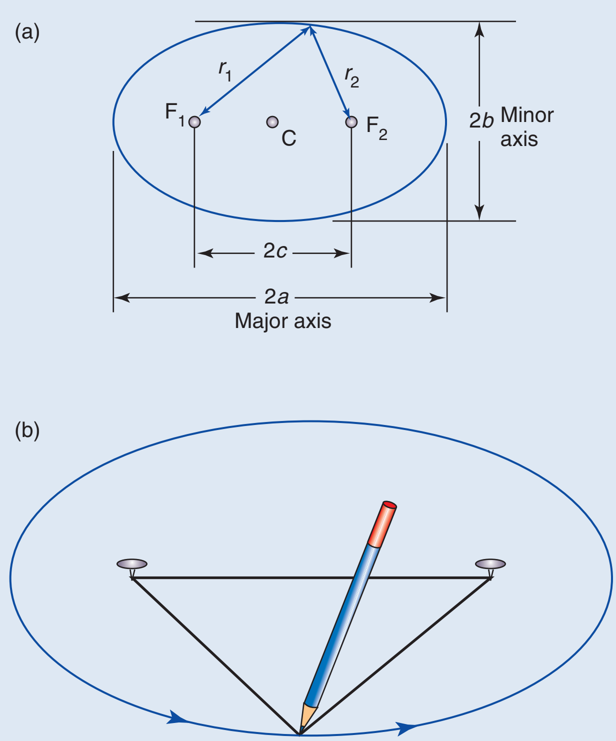


***Figure 12.23** (a) Dimensions of an ellipse (b) A method for drawing an ellipse*

For any point along an ellipse, the sum of the lengths, r_1 and r_2, equals the length of the major axis, shown as $2a$. This fact can be used to draw an ellipse using the method shown in figure 12.23b. If a string is tied into a loop and the loop is wrapped around two drawing pins stuck into a board, then a pen pushed around the inside of the loop will describe an ellipse. A relationship can be derived between the string length, s, the distance between the pins, $2c$, and the eccentricity, e, as follows:

$$\text{string length, } s = \text{major axis} + \text{pin separation}$$
$$= 2a + 2c$$
$$2a = s - 2c.$$
$$\text{Eccentricity, } e = \frac{2c}{2a}$$
$$e = \frac{2c}{s - 2c}$$

This can be rearranged to give:

$$\text{pin separation } 2c = \frac{es}{1 - e}.$$

This last expression can be used to draw ellipses of specific eccentricity using the string method, by calculating how far apart to place the pins.

Method

(a) Your string should be 15–30 cm in length. A larger dimension makes for an easier exercise, but a smaller dimension fits more easily onto a sheet of paper. Tie the two ends of your string together to make a loop. Measure the length of the loop and double this to find the string length, s, and then record this information. Push the two drawing pins into the board some distance apart and wrap the string around them as shown in figure 12.23b. Put your pen inside the loop and draw the ellipse. Measure the distance between the pins, $2c$, and the length of the major axis, $2a$, of the ellipse you have just drawn. Calculate the eccentricity of your ellipse using the two methods shown — one is the eccentricity definition $\frac{c}{a}$ and the other is the string-length equation. Set your results out in the manner shown on page 271.

(b) The following table shows the eccentricities of the eight planets in our solar system. Use the string-length equation to calculate the required pin separations to draw each of these orbits using your length of string. Tabulate this information in an extra column as suggested by the table. Once you have determined this information, go ahead and draw each of the orbits.

Eccentricities of the planetary orbits (shown to two significant figures)

PLANET	ECCENTRICITY	PIN SEPARATION (cm)
Mercury	0.21	
Venus	0.01	
Earth	0.02	
Mars	0.09	
Jupiter	0.05	
Saturn	0.06	
Uranus	0.05	
Neptune	0.01	

Results

(a) String length, s = ____________ cm

Pin separation, $2c$ = ____________ cm

Major axis, $2a$ = ____________ cm

(i) Eccentricity, $e = \frac{c}{a} = \frac{2c}{2a} =$

(ii) Eccentricity, $e = \frac{2c}{s - 2c} =$

(b) Complete the drawings of the planetary orbits.

Questions

1. Did your two calculations for eccentricity in part (a) agree?
2. Which planet has the most eccentric orbit?
3. How close to circles do the orbits of the planets appear to you?
4. Are you surprised that it took two thousand years of observation before anyone realised that the planets move in ellipses? Explain your answer.
5. What advantage did Johannes Kepler have over previous astronomers that allowed him to make this inference?

PRACTICAL ACTIVITIES

12.2 THE HUBBLE CONSTANT

Aim

To estimate a value for the Hubble constant

Theory

The Hubble constant, H_0, comes from Hubble's Law: $v = H_0 D$. It is the increase in the velocity of galaxies receding from us, for every extra megaparsec of distance they are away from us. Stated another way, if you are observing two galaxies and one of them is one megaparsec further away than the other, then the furthest one will appear to be moving away H_0 km s^{-1} faster than the closer galaxy. This means it is a measure of how fast the universe is currently expanding.

Astronomers have found a variety of ways to try to measure H_0. During the 1970s, 1980s and 1990s values were derived that consistently fell within the range between 50 and 100 km s^{-1} Mpc^{-1} (recall that one megaparsec (Mpc) equals 31 billion billion kilometres), with a variety of acknowledged errors. In 1999, the Hubble Space Telescope (HST) Key Project team began trying to find a consensus among all these values, and in 2001 released their final result: $H_0 = 72 \pm 8$ km s^{-1} Mpc^{-1}. In other words they are confident the value for H_0 lies between 64 and 80. Since that time the WMAP space probe survey of the cosmic background radiation has allowed the determination of a number of cosmological parameters — their published value for H_0 is 73.5 ± 3.2, and they have acknowledged a consensus value of 70.8 ± 1.6 km s^{-1} Mpc^{-1}.

Why is this important? It is an indicator of the age of the universe. If it is assumed that the expansion has been uniform, so that the value of H_0 hasn't changed, then the age of the universe can easily be calculated as follows:

Hubble's Law: $v = H_0 D$, so that $D = \frac{v}{H_0}$

The standard distance equation is: $D = vt$

Combining these two equations: $vt = \frac{v}{H_0}$

So that the age of the universe is given by: $t_H = \frac{1}{H_0}$

where

t_H = 'Hubble time', a rough estimate of the age of the universe.

This relationship can also be expressed as:

$$H_0 \, t_H = 1$$

This is a simple equation to use but unfortunately the assumptions are not valid — to begin with we know that the expansion of the universe has not been uniform. For some decades now the Big Bang Theory has incorporated the idea of a rapid initial expansion followed by a slow down. Further, we now know that currently the expansion of the universe is accelerating. Therefore, initially H_0 was very high, then low and is now increasing again.

As a result, the relationship between H_0 and the age of the universe t_0 becomes more complex:

$H_0 t_0 =$ a function of the density and composition of the universe as well as the acceleration of its expansion ≈ 1

Using this equation and their value for H_0, the HST team estimated the age of the universe as between 12 and 14 billion years old. The WMAP is more confident, however. Using their own data they have been able to measure the density, composition and acceleration of the universe (see page 264) as well as H_0, and have calculated that the universe is 13.7 billion years old ±1%. These calculations are based on the universe model that has had most consensus since 1998 — the Lambda CDM model, which is a spatially-flat universe made of atoms, cold dark matter and dark energy.

Method

Using the Hubble Space Telescope, the project team used a variety of methods — such as using a type of pulsating star called a Cepheid variable — to measure the distances to galaxies much further away than can be achieved from the ground. The relative velocity of recession of each galaxy was also measured using their Doppler red shift. Some of the results are shown in the table below.

DISTANCE, *D* (Mpc)	VELOCITY, *v* (km s^{-1})
15	1 100
97	6 700
32	3 000
145	10 700
50	3 100
122	9 900
58	4 300
91	5 300
120	9 000
93	7 500
158	8 900
64	5 300
145	9 600
61	3 300
103	5 100
46	5 000
34	1 800
185	9500
15	1 700
20	1 200

1. Draw a graph of this data, with distance along the horizontal axis and velocity along the vertical axis.
2. Draw a line of best fit (it should pass through the origin) and evaluate the gradient. This will be your estimate of the value of the Hubble constant, H_0.
3. Draw two more lines (again through the origin) that represent an upper and a lower boundary of the data. The gradients of these lines represent the maximum and minimum values of H_0 that can fit the data set.

Results

Maximum possible value of H_0 from data = ______

Best fit value of H_0 = ____________

Minimum possible value of H_0 from data = _______

Hence, from this data, H_0 can best be stated as having a value of ___________ with an uncertainty of __________ %.

Questions

1. What have been the accepted upper and lower limits for H_0 for the last few decades?
2. Complete the following table using the information provided in the Theory and Results sections.

	HST TEAM	WMAP DATA	WMAP CONSENSUS	YOUR RESULTS
H_0 value				
Stated uncertainty				
H_0 upper limit				
H_0 lower limit				

3. If $H_0 t_0 = 1$, calculate the age of the universe (Hubble time in this case) using your value for H_0. Note that you will need to convert the distance units, kilometre and megaparsec, into metres first.
4. Repeat this calculation using your upper and lower limits for H_0, thereby determining an upper and lower limit for t_0.
5. Compare your calculated values for the age of the universe with those published by the HST and WMAP teams.

CHAPTER 13 STAR LIGHT, STAR BRIGHT

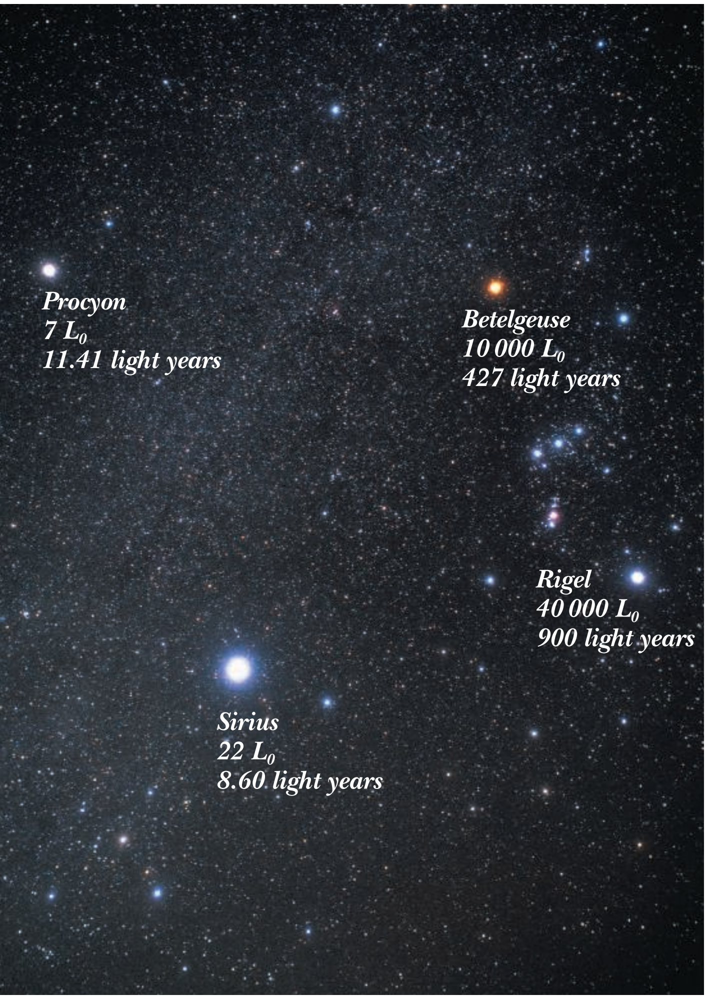


Figure 13.1 *A portion of the night sky near the constellation of Orion. Four of the brightest stars in the sky can be seen here — Sirius, Betelgeuse, Rigel and Procyon.*

Remember

Before beginning this chapter, you should be able to:

- describe the features and locations of protons, neutrons and electrons in the atom
- describe some changes that are likely to take place during the life of a star.

Key content

At the end of this chapter you should be able to:

- use the inverse square law to quantitatively relate the brightness of a star to its luminosity and distance from the observer
- use Wien's Law to define the relationship between the temperature of a body and the dominant wavelength of the radiation emitted from that body
- relate the surface temperature of a star to its colour
- describe a Hertzsprung–Russell (H–R) diagram as the graph of a star's luminosity against its colour or temperature
- identify star groups on an H–R diagram, including main sequence, red giants and white dwarfs
- identify energy sources characteristic of each of the main star groups on an H–R diagram.

Note that the colours of the stars in figure 13.1, as in all star photographs, appear slightly different to those normally ascribed to them due to the specific colour sensitivities of photographic film.

If you look into the early evening sky in summer you can see the constellation of Orion, the hunter. In and near this constellation are found four of the brightest stars in the sky — Betelgeuse (pronounced BEET-el-jooz), Rigel (RYE-jel), Procyon (PRO-see-on) and Sirius.

If you look at these stars, there are two features of each that are quickly noticed — their brightness and their colour. Betelgeuse, Rigel and Procyon are all about the same brightness, while Sirius is slightly brighter (in fact, the brightest star in the night sky). Also, Sirius and Rigel are both a blue-white colour, while Procyon is yellow and Betelgeuse is red. Using just these two features — brightness and colour — much can be learned about a star, and stars in general, relating to temperature, distance, star type and energy source. These four stars will be used as examples to illustrate these points in the early parts of this chapter.

13.1 A STAR'S LUMINOSITY AND BRIGHTNESS

The luminosity of a star is the total energy radiated by a star per second.

The brightness of a star is the intensity of light as seen some distance away from it. It is the energy received per square metre per second.

Luminosity is the total energy radiated by an object (such as a lamp or a star) per second. This can also be called power output and its SI units are joules per second, or watts (W). A light globe may have a power rating of 100 W, while the Sun's power output, or luminosity, is estimated to be approximately 3.83×10^{26} W. This value is designated L_0 and is often used as a unit to express the luminosity of other stars. In figure 13.1 the luminosity of Procyon is noted as 7 L_0 .

The **brightness** of a radiant object (such as a lamp or star) is the intensity of light as seen some distance away from it. It is the energy received per square metre per second.

Referring back to figure 13.1, Procyon and Betelgeuse are of almost equal brightness, yet Betelgeuse (10 000 L_0) is more than one thousand times more luminous than Procyon (7 L_0). How is it possible that they can both appear of equal brightness to us? The answer is that Betelgeuse is about forty times further away from us than Procyon, and this distance causes an apparent dimming of the light we see. This is similar to looking at streetlights at night — the closest light can be quite bright, but as you look further into the distance, the streetlights appear dimmer and dimmer. Figure 13.2 serves to highlight the relationships between the factors affecting the brightness of a star.

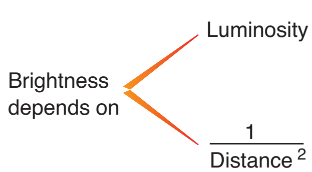


Figure 13.2 *Brightness and luminosity are related.*

The inverse square law

To understand this relationship better, consider the construction shown in figure 13.3. This diagram shows a star surrounded by an imaginary sphere. The star gives off radiant energy in all directions, and the sum of all the radiation given off in one second is the star's luminosity. The radiation spreads out uniformly and penetrates the whole surface of the sphere. Therefore, the amount of radiant energy per square metre per second received at the surface of the sphere (the brightness) is given by:

$$\text{brightness} = \frac{\text{luminosity}}{\text{surface area of the sphere}}.$$

Since the surface area of a sphere $= 4\pi r^2$

$$\text{then, brightness} = \frac{\text{luminosity}}{4\pi r^2}.$$

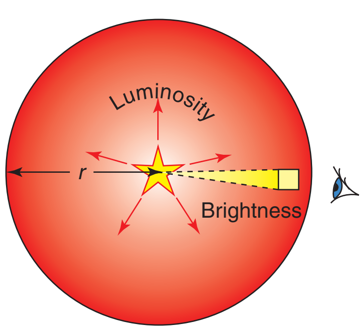


Figure 13.3 *The total radiant energy from a star is spread over the surface of an imaginary sphere surrounding it.*

PRACTICAL ACTIVITIES

13.1

The inverse square law

This relationship describes how the brightness of a star, a distance r from an observer, depends on the luminosity and the distance. In particular, notice that the brightness is inversely proportional to the square of the distance. This can be written as follows:

$$\text{brightness} \propto \frac{1}{r^2}.$$

This relationship means that if the distance, r, were to double, then the brightness would reduce to one-quarter of its previous value.

$$\begin{aligned}\text{New brightness} &\propto \frac{1}{(2r)^2}\\ &\propto \frac{1}{4r^2}\\ &\propto \frac{1}{4}\times\frac{1}{r^2} = \frac{1}{4}\ \text{of previous brightness}\end{aligned}$$

SAMPLE PROBLEM 13.1

Brightness of a star

Use the data provided in figure 13.1 to show that Procyon and Betelgeuse have almost the same brightness.

SOLUTION

For Procyon:

$$\begin{aligned}\text{brightness} &= \frac{\text{luminosity}}{4\pi r^2}\\ &= \frac{7\text{L}_0}{4\pi(11.41)^2} = 0.004\ \text{L}_0\ \text{watts per square light year.}\end{aligned}$$

To express this in SI units we must note that:

$$\begin{aligned}11.41\ \text{light years} &= 11.41\times 9.46\times 10^{15}\ \text{m}\\ &= 1.08\times 10^{17}\ \text{m}\end{aligned}$$

$$\begin{aligned}\text{and therefore,}\ \ \text{brightness} &= \frac{\text{luminosity}}{4\pi r^2}\\ &= \frac{7(3.83\times 10^{26})}{4\pi(1.08\times 10^{17})^2} \approx 2\times 10^{-8}\ \text{W m}^{-2}.\end{aligned}$$

For Betelgeuse:

$$\begin{aligned}\text{brightness} &= \frac{\text{luminosity}}{4\pi r^2}\\ &= \frac{10\,000\ \text{L}_0}{4\pi(427)^2} = 0.004\ \text{L}_0\ \text{watts per square light year.}\end{aligned}$$

Therefore, both stars have a brightness of 0.004 L_0 watts per square light year.

SAMPLE PROBLEM 13.2

Brightness of a star relative to distance

How would the brightness of Procyon change if its distance were three times its current value; that is, if it were at a distance of 34.23 light years?

SOLUTION

The simplest path to a solution in this case is to use the inverse square law — if the distance is increased by a factor of three, then the brightness will be reduced by a factor of 3^2 (nine).

$$\begin{aligned}\text{Therefore, new brightness} &= \frac{\text{old brightness}}{9}\\ &= \frac{0.004\,\text{L}_0}{9} \approx 0.00044\ \text{L}_0\ \text{watts per square light year.}\end{aligned}$$

Let us check this with a calculation using the brightness expression:

$$\text{brightness} = \frac{\text{luminosity}}{4\pi r^2}$$

$$= \frac{7L_0}{4\pi(34.23)^2} \approx 0.00047\ L_0 \text{ watts per square light year.}$$

The difference between the answers is due to a rounding off error.

It was mentioned earlier that the luminosity of the Sun has a value of approximately 3.83×10^{26} W. This value was determined using the brightness/luminosity relationship. Detectors aboard satellites were used to measure the amount of radiant energy per square metre per second reaching us from the Sun; that is, the brightness of the Sun. In addition, the distance of the Earth from the Sun is well known at any point around its slightly elliptical orbit. Using these two pieces of information, the luminosity of the Sun could be calculated.

13.2 TEMPERATURE AND COLOUR

When looking at stars such as those shown in figure 13.1, the next feature you may notice after brightness is the colour of a star. You will notice this especially if you are using binoculars or a telescope to view them. Of the four brightest stars in figure 13.1, two are a blue-white colour, another is yellow and the fourth is red. What factors have led to these different colours? The answer is that temperature is mostly responsible; but to appreciate why, you will need to understand the basics of black-body radiation.

A **black body** is one that absorbs all radiation falling upon it. When it becomes hotter than its surroundings, it begins to radiate electromagnetic energy of its own.

Black-body radiation is the electromagnetic radiation emitted by a black body.

A **black body** is one that absorbs all radiation falling upon it. When it becomes hotter than its surroundings, it begins to radiate electromagnetic energy of its own. This is known as **black-body radiation**. The radiation is distributed continuously but not evenly across the various wavelengths of the electromagnetic spectrum. The way that it is distributed depends on the temperature of the black body (see figure 13.4).

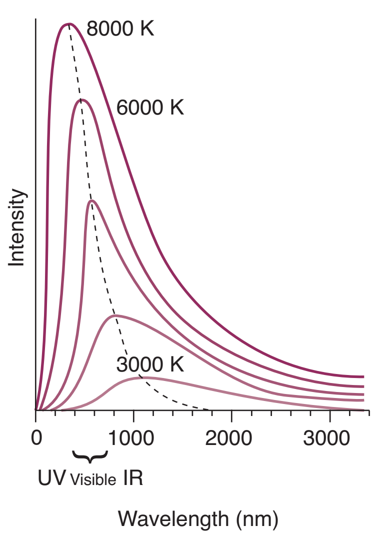


***Figure 13.4** Black-body radiation curves. As a black body becomes hotter the peak of the curve shifts to shorter (bluer) wavelengths.*

Each of the curves in figure 13.4 is specific to a particular temperature, and each has a peak intensity that corresponds to a particular wavelength. As the temperature increases you can see that the peak moves towards the shorter wavelengths. At lower temperatures, the radiation lies mostly in the non-visible infra-red region; but as the temperature increases, the peak moves into the visible spectrum. At higher temperatures, the peak has moved out to the ultraviolet region. This relationship is known as Wien's Law and can be written as:

$$\lambda_{max} T = W$$

where

λ_{max} = wavelength of maximum output or dominant wavelength (m)
T = temperature (K)
W = a constant = 2.9×10^{-3} m K.

Although a black body is a hypothetical concept, a star's radiation output closely matches the theoretical black-body radiation curves, at least in and near the visible spectrum, which determines its colour. Refer again to figure 13.4 and consider how the colour we see coming from a

13.2

Relating colour to temperature

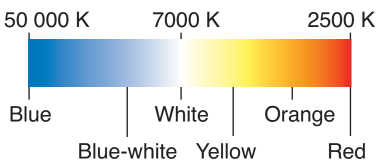


Figure 13.5 *Colour and surface temperature are directly related.*

Weblink:
Black-body radiation

star is affected by the shifting black-body radiation curves. In particular, look at the balance of radiation within the visible spectrum. Note the following points:

- A cooler star, with a surface temperature of perhaps 3000 K or 4000 K, produces most of its visible radiation at the longer wavelength, or red, end of the spectrum. It will therefore appear red.
- As the temperature increases, the wavelengths peak in the yellow part of the spectrum, and a star will appear yellow.
- At slightly higher temperatures, the distribution of the radiation is more even and, as a result, a star will appear white.
- A hotter star, with a surface temperature of 8000 K or more, produces most of its visible radiation at the shorter wavelength, or blue, end of the spectrum. Therefore, it appears blue.

Therefore, there is a correlation between a star's surface temperature and its colour, as shown in figure 13.5. Hence, referring back to figure 13.1, we can now say that red Betelgeuse is the coolest of the four bright stars shown; Procyon, being yellow, has a surface temperature of about 6000 K (similar to our Sun); Sirius and Rigel are the warmest of the four, because they are a blue-white colour.

Wien's Law allows us a more accurate way to measure a star's temperature. A device known as a spectrophotometer can be used to analyse a star's light, to measure the wavelength of maximum output. Once this is known, the star's surface temperature can be calculated directly using Wien's Law. As accurate as this may sound, even better determinations of a star's temperature can be made by analysing the individual wavelengths present in a star's spectrum. These techniques are covered in greater depth in the 'Astrophysics' module in the HSC course.

13.3 THE HERTZSPRUNG–RUSSELL DIAGRAM

In 1911, the Danish astronomer Ejnar Hertzsprung devised a diagrammatic way of representing the properties of a star. The technique was independently developed in 1913 by the American astronomer Henry Norris-Russell. Today the method is referred to as a **Hertzsprung–Russell diagram**, or simply an H–R diagram. It is one of the most significant analytical tools for the study of stars and their evolution.

A Hertzsprung–Russell diagram is a graph of a star's luminosity (as the vertical axis) plotted against its temperature or colour.

An H–R diagram is a graph of a star's luminosity (on the vertical axis) plotted against its temperature or colour. However, the axes are not linear. The vertical luminosity scale extends in value from perhaps 10^{-4} L_0 up to a value of 10^6 L_0, with equal units indicating an increase by a factor of ten. The horizontal temperature scale extends from 2500 K at the right to perhaps 40 000 K at the left, with equal units indicating a doubling of temperature. A single star plots as a single point on such a graph. Figure 13.6 shows an H–R diagram with the plotted positions of the four stars from figure 13.1, as well as the Sun. It is clear from this graph alone that there are different types of stars.

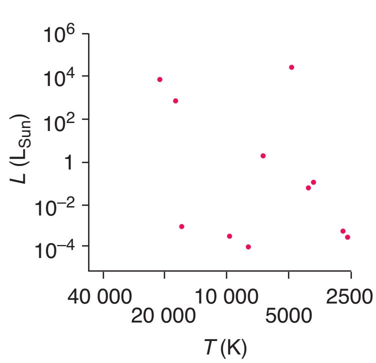


Figure 13.6 *An H–R diagram with the positions of just a few stars plotted*

However, the power of the H–R diagram lies in its ability to show whole groups of stars, so that the relationships between them can be seen. Figure 13.7a on page 278 shows an H–R diagram for the nearest and brightest stars. Plotting a selection of stars such as this provides an assortment of types, as the graph shows. The stars tend to fall into a number of discrete groups, which are clearly shown in figure 13.7b.

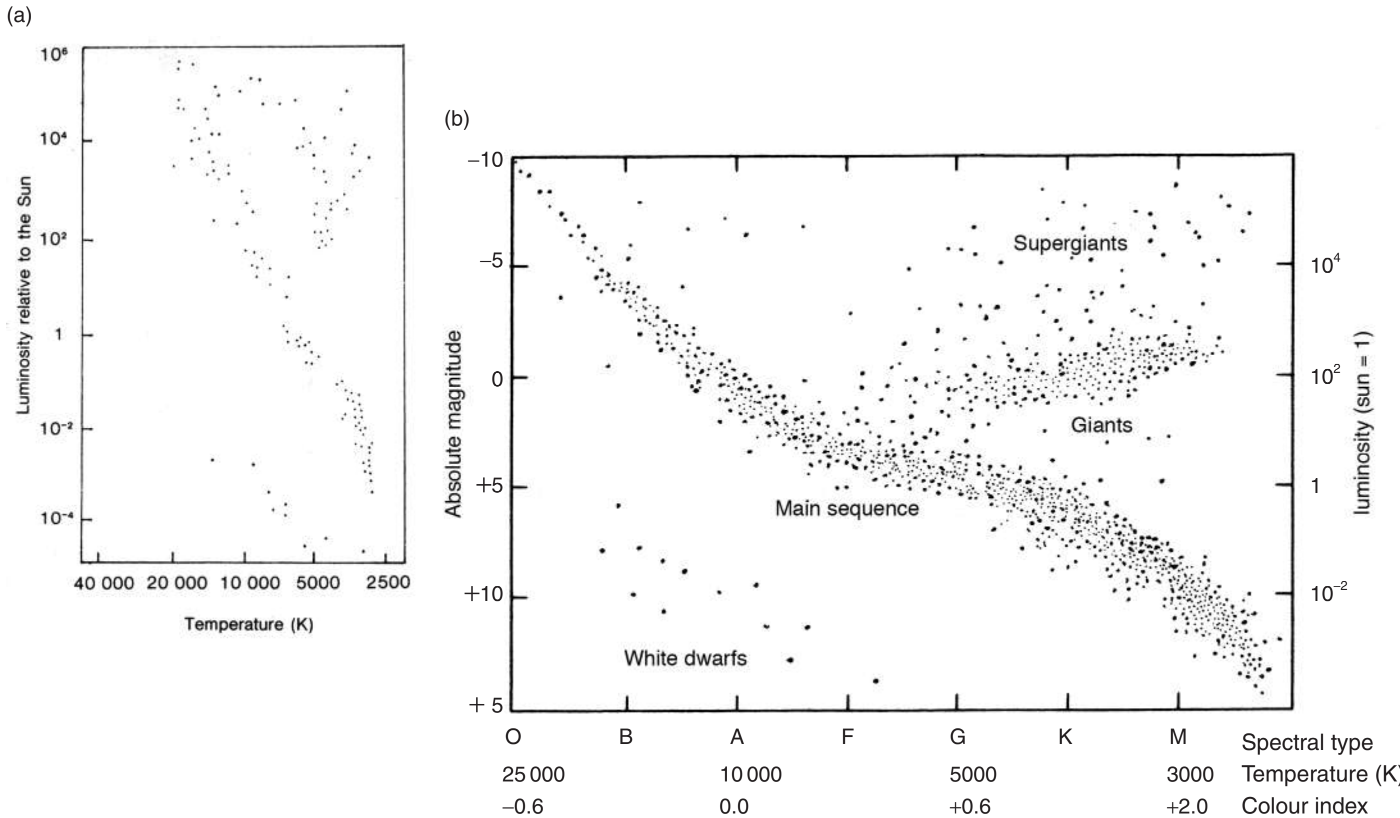


Figure 13.7 *(a) An H–R diagram for the nearest and brightest stars and (b) the star groups revealed*

Identifiable star groups

The basic star groups, revealed on an H–R diagram (see figure 13.7b), are:

The main sequence is a diagonal band from the upper left corner of an H–R diagram to the lower right corner of the diagram. Most stars are in this group.

- the **main sequence**. Most stars line up along a slightly curved diagonal band from the upper left corner of the diagram to the lower right corner. This is referred to as the main sequence, and it represents a variety of star masses. Through a close study of binary (or double) stars, it has been learned that more luminous main sequence stars are also more massive. Hence, the upper end of the main sequence represents hot stars that are as much as one million times more luminous and one hundred times more massive than the Sun. The lower end represents cool stars that are as little as one-thousandth the luminosity and one-hundredth the mass of the Sun.

The red giants is the group of stars in the upper right corner of the H–R diagram. These are cool, giant stars.

- the **red giants**. This is the group of stars in the upper right corner of the H–R diagram. They are both cool and luminous. A star's luminosity is very sensitive to its temperature. Therefore, the high luminosity of these stars can only be due to their extraordinary size; that is, they are giant stars. There appear to be two groups of these stars — the red giants (red indicates their cool surface temperature) and the more luminous, and therefore even larger, supergiants.

The white dwarfs is the group close to the lower left corner of the H–R diagram. These are exceptionally small, hot stars.

- the **white dwarfs**. This group lies below the main sequence, close to the lower left corner of the H–R diagram. These stars are white-hot yet they have very low luminosities. We would normally expect stars of this temperature to be ten thousand times more luminous than these stars. As with red giants, it is the stars' unusual size that results in their unusual luminosity, except that in this case they are exceptionally small. They are called the white dwarfs.

13.4 ENERGY SOURCES WITHIN STAR GROUPS

The size of a star is the result of a balance between two forces, as shown in figure 13.8. One is the force of gravity, trying to compress the star. This is the force that originally caused the star to form from a gas cloud and compressed it sufficiently for it to heat up. For this reason, gravity can be thought of as a source of energy for a star.

Working against the force of gravity is the star's radiation pressure, an outwards force resulting from the nuclear reactions within the star. Without this radiation pressure the star would simply be crushed by its own gravity. The radiation pressure varies between red giants, white dwarfs and main sequence stars. This results in the differences in star sizes. The different radiation pressures are due to different nuclear reactions within the cores of the stars.

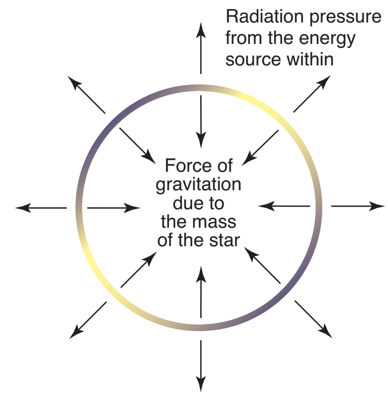


***Figure 13.8** The two forces that determine the size of a star*

Nuclear fusion in the star groups

eBookplus
Weblink:
Nuclear fusion in the Sun

In 1938, Hans Bethe and Carl von Weizsächer independently proposed that the nuclear fusion of hydrogen to helium would provide sufficient fuel for a star. They then went further and provided a reaction mechanism that could achieve this. Their mechanism has been called the 'carbon–nitrogen–oxygen cycle', because these elements act as catalysts, helping the reaction to occur. Today it is recognised that this is not the only mechanism for the fusion of hydrogen to helium inside a star. Another, called the 'proton–proton chain', involves no catalysts and can occur at lower temperatures. Whichever mechanism is followed, the overall reaction is the same. As shown in the following reaction, four hydrogen nuclei join to form a single helium nucleus, releasing two positrons (anti-electrons), two neutrinos and some energy.

$$4\,{}^{1}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He} + 2e^{+} + 2\nu + 4.3 \times 10^{-12}\ \mathrm{J}$$

This is the reaction that occurs within the core of main sequence stars, which have a large supply of hydrogen. However, as the reaction proceeds, the hydrogen supply dwindles and a supply of helium is built up. The next energy-producing reaction that can occur is the fusion of helium to carbon. This reaction is known as the 'triple-alpha reaction' (because a helium nucleus is an alpha particle), and is shown in the following reaction. In this reaction, three helium nuclei fuse to form a single carbon nucleus, releasing gamma radiation as well as some energy.

$$3\,{}^{4}_{2}\mathrm{He} \rightarrow {}^{12}_{6}\mathrm{C} + \gamma\ \text{radiation} + 1.2 \times 10^{-12}\ \mathrm{J}$$

This reaction is recognised as the reaction that occurs in the core of red giants, while hydrogen fusion continues in a shell around the core. In more massive stars, further energy-producing fusions can occur. For instance, carbon can fuse with helium to form oxygen.

$${}^{12}_{6}\mathrm{C} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{16}_{8}\mathrm{O} + \gamma$$

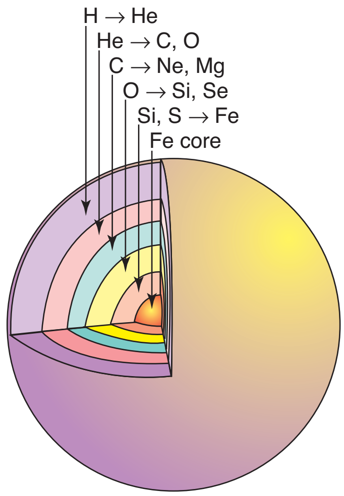


Figure 13.9 ***A supergiant can contain many energy sources within a series of shells.***

In supergiants, the oxygen can fuse to heavier elements such as silicon or sulfur. These can fuse further, forming elements as heavy as iron, which is the heaviest element that can be formed in an energy-producing reaction. A system of multiple shells, each with a different energy source, can exist, as shown in figure 13.9.

White dwarfs are different to both main sequence stars and giant stars, in that they have no nuclear fusions occurring within them. As we shall see in the following section, they are collapsed star corpses, with no nuclear fuel left to burn. With no energy source remaining, white dwarfs simply radiate their heat into space and cool down to become black dwarfs.

Table 13.1 summarises these reactions.

Table 13.1

STAR GROUP	ENERGY-PRODUCING REACTIONS
Main sequence	Nuclear fusion of H to He in core
Red giants	Nuclear fusion of He to C in core, with H fusion continuing in shell
Supergiants	Multiple nuclear fusions possible in shells, forming elements up to iron in core
White dwarfs	No energy-producing reactions occurring

Using the knowledge of the sequence of energy-producing nuclear fusion reactions, along with other evidence, it is possible to construct a description of the life of a typical star. The description that follows can also be plotted on an H–R diagram, as shown in figure 13.10. This plot shows the path that would be followed by a star of approximately one solar mass.

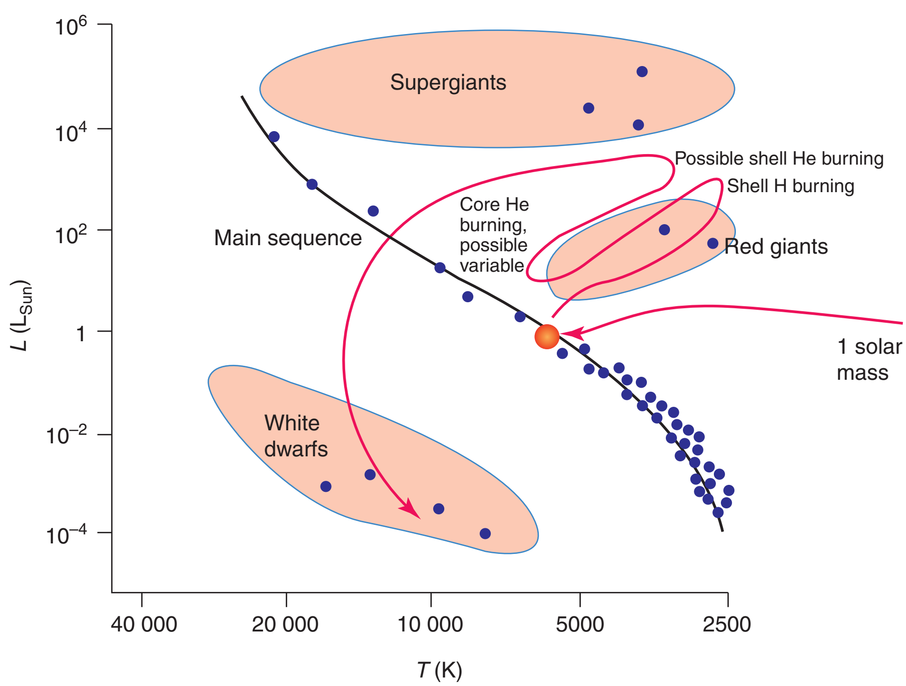


Figure 13.10 ***An H–R diagram showing the path followed by a star of approximately 1 solar mass during its lifetime***

The protostar stage

A star begins its life as a contracting gas cloud. This cloud is predominantly hydrogen, in atomic as well as molecular form, with some helium. The cloud may contain a small percentage of larger organic molecules such as carbon monoxide, methane and ethanol. If other stars have existed and expired nearby, then the cloud will be enriched with heavier elements, possibly as heavy as lead. If the expired star was sufficiently massive that it finished with a supernova explosion, then even heavier elements such as uranium may be present. This was the case with our Sun, as indicated by the presence of these heavy elements on Earth.

The contracting cloud heats up as its contracts and can be called a 'protostar'. If the protostar is sufficiently massive (approximately 0.01 solar masses), then the core will heat up enough for the fusion of hydrogen to begin. If it is too big (greater than approximately 100 solar masses), then once fusion begins the radiation pressure will be so great that the protostar is immediately blown apart.

The main-sequence stage

Once the fusion of hydrogen begins in its core, the protostar becomes a main sequence star and will spend most of its life in this form. The position a star takes up on the H–R diagram depends on its mass — higher mass stars are positioned further up the main sequence.

With hydrogen fusing in its core, a main-sequence star produces helium, which settles to the centre of the star because it is denser than hydrogen. Eventually, the star has only a dwindling fuel supply, insufficient to continue to support the star, which is ready to evolve to its next stage.

The red-giant stage

The next stage involves the radiation pressure (produced by the fusion) petering out and the star's core shrinking. The remaining hydrogen fusion shifts to a shell that expands greatly in size, cooling as it becomes less dense. Meanwhile, the core collapses due to gravitation. This causes a sudden increase in pressure and temperature that starts the fusion of helium to carbon.

If the star has less than approximately one-fifth of the mass of the Sun then it is doubtful that it will ever achieve the temperature required to begin helium fusion. If the star is quite massive then, as helium fusion begins to wane, further fusions can start, leading to a multilayer structure.

The white-dwarf stage

Eventually the helium-fusing core of a red giant runs out of fuel. More massive stars, engaging in post-helium burning reactions, also run out of their fuel sources. In each case, the star is facing the final stage of its life. As the star contracts yet again due to the ever-present gravitational force, it sheds its outer layers and so loses mass. The core that remains is a hot and dense star corpse, no longer performing fusion reactions and slowly cooling down.

A star of approximately 1 solar mass loses its outer layers gently, shedding them as a sphere-shaped nebula, known as a planetary nebula. The corpse that remains is about the size of the Earth and is called a white dwarf. This corpse eventually cools down to be a black dwarf.

A star of greater mass loses its outer layers in a much more violent manner — an enormous explosion called a supernova. The remaining stellar corpse is much denser than a white dwarf — either a neutron star or a black hole.

The various stages of a star's life, as well as its ultimate end point, are most dependent on its mass. As shown in figure 13.11, the mass of the original cloud determines whether it will form a star, and where on the main sequence it becomes located. The mass of the star determines how long it will remain on the main sequence, what type of giant it eventually becomes and how the star sheds it outer layers. The mass of the core of the star determines what form the stellar corpse will take.

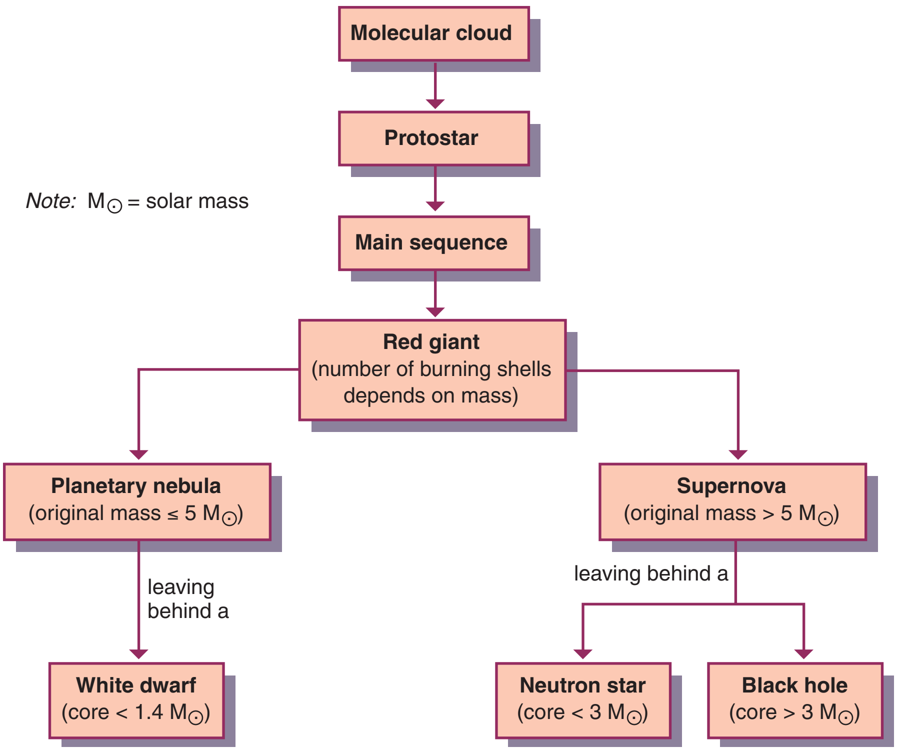


Figure 13.11 *Stellar evolution depends on mass*

SUMMARY

- Luminosity is the total energy radiated by a star per second. It is measured in watts.
- The brightness of a star is the intensity of light seen some distance away from it. It is the energy received per square metre per second.
- Brightness and luminosity are related by the inverse square law:

$$\text{brightness} = \frac{\text{luminosity}}{4\pi r^2}.$$

- A black body is one that absorbs all radiation falling on it. When it becomes hotter than its surroundings, it begins to radiate electromagnetic energy of its own, the distribution of which depends on the black body's temperature.
- A star's energy output approximates that of a black body. Therefore, a star's colour depends on its temperature. The series of observable colours from cool to hot is: red, yellow, white, blue-white, blue.
- A Hertzsprung–Russell (H–R) diagram is a graph of a star's luminosity (as the vertical axis) plotted against its temperature, or colour.
- The main star groups apparent on an H–R diagram are the main sequence, red giants and white dwarfs.
- A newly formed main sequence star is mostly hydrogen, which it fuses into helium. Later, as a red giant, the star fuses this helium into carbon. When all fusion reactions cease, the non-fusing core forms a white dwarf. With no remaining energy source, the white dwarf eventually cools down to form a cold stellar corpse.

QUESTIONS

1. (a) Define luminosity.
   (b) Define brightness.
   (c) Describe the relationship between these two concepts.
   (d) Describe the inverse square law as it relates to the brightness of a star.
2. Proxima Centauri is slightly more than 4 light years away. Imagine you are on a journey to this star. Compared to the brightness of the star at the beginning of the journey, how bright would the star appear when you are (a) half, and (b) three-quarters of the distance to the star?
3. (a) If you observe two stars in the night sky and one is brighter than the other, does this mean that it is more luminous than the other? Explain your answer.
   (b) The stars in a star cluster are all approximately the same distance from us. If you compare two stars in a cluster and notice that one is brighter than the other, does this mean that it is necessarily more luminous? Explain.
4. You observe two stars in the sky and notice that they look very similar. After research, you discover that one of the stars is four times further away than the other. Compare the brightness and luminosity of the two stars.
5. Two of the brightest stars in the sky are Sirius and Rigel. Sirius has a luminosity of 22 L_0 and is 8.60 light years away. Rigel has a luminosity of 40 000 L_0 and is 900 light years away. Compare the brightness of these two stars.
6. Fomalhaut is a star that is 25 light years away and is 17 times more luminous than the Sun. Bellatrix is 243 light years away and 1000 times the luminosity of the Sun. Determine which of these two stars appears the brighter in the night sky.
7. (a) Describe the concept of a black body.
   (b) Explain the relevance of this concept to stars.
8. (a) Construct a sample black-body radiation curve. Label the vertical and horizontal axes.
   (b) Indicate the wavelength of maximum energy output. Describe how this wavelength will change if the black body becomes hotter.
9. Describe the relationship between a star's temperature and its colour.
10. Draw a vertical line several centimetres long. Above the line write 'hot' and below it write 'cool'. Mark out the area to the left of the line as a temperature scale and the area to the right of the line as the corresponding colour scale.
11. (a) Identify two variables of a star that are plotted on a Hertzsprung–Russell diagram.
    (b) Identify the correct axis for each variable.
12. Construct an H–R diagram and indicate on it the main star groups.
13. Identify the generalisation that can be made about the masses of the stars in the main sequence.

14. List each of the major star groups and provide a description of the stars found in each.
15. Copy the following table and then complete it by writing a short description of the nuclear reaction occurring at each location indicated.

STAR TYPE	CORE REACTION	SHELL REACTION
Main sequence		
Red giant		
White dwarf		

16. On the diagram you drew for question 12, draw the path followed by a star of 1 solar mass during the course of its life.
17. Explain why the reactions identified in question 15 suggest the sequence for the life of a star, as shown in your answer to question 16.

13.1 THE INVERSE SQUARE LAW

Aim

To observe the inverse square relationship between brightness and distance

Apparatus

bright lamp
data logger with light sensor or a light meter

Background information

Light is emitted from a source, such as a star or a lamp, in all directions. Rays of light at a distance from a point source are spread over the surface of a sphere of radius, r, and area, $4\pi r^2$, so that at any place upon that surface, the brightness of light is proportional to the inverse of the square of the distance: $I \propto \frac{1}{r^2}$. This is known as the inverse square law.

Method

Darken the room and turn on your lamp. The most suitable starting point will depend on the power of your lamp and the sensitivity of your sensor. Place the light sensor at a distance that produces a near 100 per cent output reading from your sensor. Move the sensor back a few centimetres at a time, recording the sensor output each time. Tabulate your results in a table similar to that shown in the results section below, and then graph the output versus $\frac{1}{r^2}$. Alternatively, if you are using a data logger, the software provided may allow you to plot the graph directly.

Results

Your results table should include at least three columns as shown here.

LIGHT SENSOR OUTPUT	DISTANCE, r	$\frac{1}{r^2}$

Questions

1. If r increases, what happens to the value of the output reading from the sensor?
2. Do the data points you have plotted appear to define a straight-line relationship between the output and $\frac{1}{r^2}$?
3. If it can be said that the output is directly proportional to $\frac{1}{r^2}$, how could we describe the relationship between output and distance, r?

13.2 RELATING COLOUR TO TEMPERATURE

Aim

To relate the colour of a hot object to its temperature

Apparatus

12 V mounted lamp
voltmeter
variable 12 V power supply
ammeter
hand spectroscope

Theory

Wien's Law for black-body radiation tells us that as an object gets hotter, the dominant wavelength of the electromagnetic radiation it emits shifts towards the blue end of the spectrum. The object in this experiment is a lamp filament. Voltage and current readings will be taken during the experiment so that the resistance of the filament can be calculated. This will be used as an indication of temperature, as the resistance of a filament is approximately proportional to its temperature.

Method

1. Set up a circuit with the power supply, ammeter and mounted lamp in series; place the voltmeter in parallel with the lamp.
2. Darken the room, set the power supply to its lowest setting and turn it on. Record the readings on the voltmeter and ammeter, then calculate the resistance of the filament using Ohm's Law.
3. Use the spectroscope to examine the dim light from the glowing filament, noting which colours are present and their relative intensities. Record this information by shading in the section of spectrum observed.
4. Repeat this process for each successively higher setting on the power supply.

Results

Your results table should resemble that shown below.

Questions

1. Does each successively higher setting on the power supply produce a higher resistance of the filament?
2. What colours are present in the spectrum on the lowest setting? Describe your impression of the filament colour, looking directly at it without the spectroscope.
3. What colours are present in the spectrum on the highest setting? Describe your impression of the filament colour this time, once again without the spectroscope.
4. Looking at your series of diagrams representing the spectra, what general change can be seen as temperature increases (aside from becoming brighter)?
5. Do your results agree with Wien's Law?

Power supply setting	Voltmeter reading, V (V)	Ammeter reading, I (A)	Resistance $R = \frac{V}{I}$ (Ω)	Spectrum,						
				V	I	B	G	Y	O	R
A										
B etc.										

PRACTICAL ACTIVITIES

CHAPTER 14 THE SUN–EARTH CONNECTION

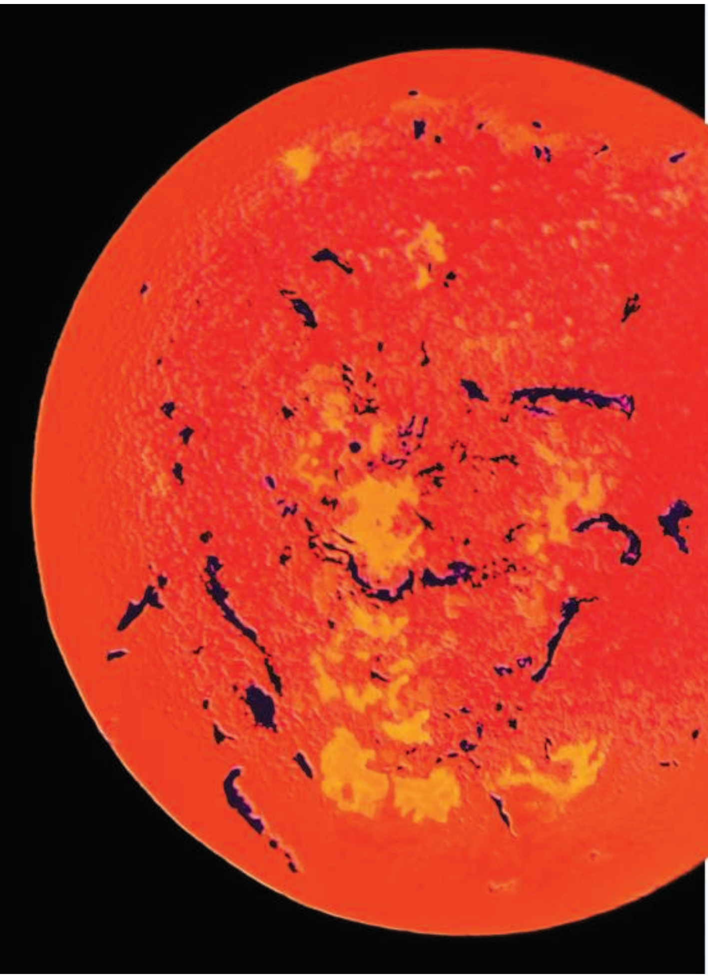


Figure 14.1 *A photograph of the Sun showing sunspots*

Remember

Before beginning this chapter, you should be able to:

- recall the features and locations of protons, neutrons and electrons in the atom
- describe the nature and creation of ions.

Key content

At the end of this chapter you should be able to:

- describe the nature of emissions from the nuclei of atoms as radiation of alpha α, beta β, and gamma γ rays
- identify that energy may be released from the nuclei of atoms
- identify that the Earth receives electromagnetic and particulate emissions from the Sun
- describe the solar winds
- describe the nature of sunspots
- outline the nature of the solar cycle and its relationship with sunspot activity and the solar wind
- assess the effects on Earth of fluctuations in the solar wind.

The Sun experiences periods of turbulent activity as well as periods of quiet. Throughout both periods, it continues to release a reasonably steady stream of sunlight to warm our planet.

During the years 1645 to 1715, barely a blemish (an indication of magnetic turbulence) was observed on the surface of the Sun. This period was known as the 'Maunder minimum'. The Sun burned with a quiet steadiness, with very little turbulence or sunspots — such as those in figure 14.1 — to be seen. During the same period, Europe experienced what has become known as 'the little ice age'. Otherwise ice-free rivers froze over and fields lay snow-covered year round. While a link between solar events and variations in the Earth's weather has not been proven, it is a field of ongoing research.

However, space weather can have a significant impact on the Earth in a variety of other ways. Space weather refers to changes in the Sun–Earth environment. It is dominated by electromagnetic and ionic emissions from the Sun, which can fluctuate wildly during periods of solar activity. The Earth's magnetic field has some linkage to the Sun's own extended magnetic field, and this allows turbulent space weather to cause geomagnetic storms on Earth, with a variety of troublesome effects.

14.1 NUCLEAR RADIATION

PRACTICAL ACTIVITIES

14.1

Radiation penetration

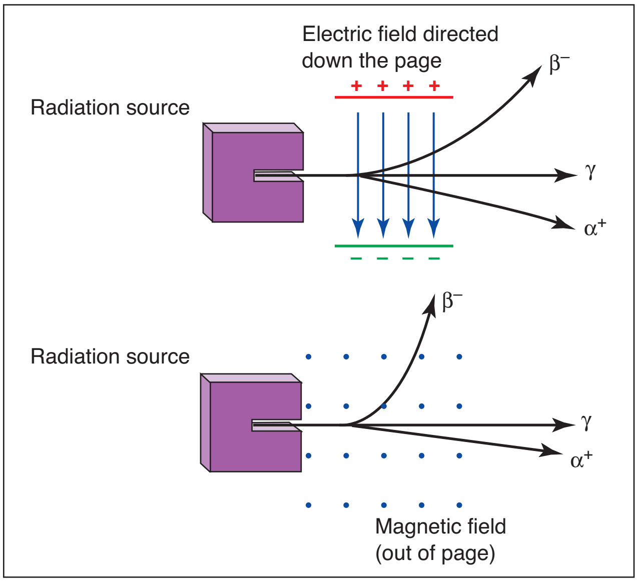
The three types of radiation are alpha α (helium nuclei), beta β (electrons) and gamma γ (high-frequency electromagnetic radiation).

Nuclear radiation is often discussed in the news media with regard to nuclear reactors, power stations and bombs — but what exactly is it? Henri Becquerel made the first observations of radioactivity in 1896 when experimenting with uranium. Soon afterwards, Marie Curie and her husband Pierre Curie realised that thorium was also radioactive, and then went on to discover two more radioactive elements, radium and polonium (although Pierre died before this work was completed). By 1906, approximately twenty radioactive elements were known. All of these elements produced some sort of emission that could be detected by its properties. These properties included the fact that they affected photographic film and that they discharged electroscopes (devices that hold electrostatic charges).

Through the work of Ernest Rutherford and many other physicists in the first decade of the twentieth century, it was realised that there are three types of radiation — **alpha α**, **beta β** and **gamma γ**. Each has a different charge, ionising and penetrating ability. The results of many experiments to determine the nature of these radiations are summarised in table 14.1.

Table 14.1 **Types of radiation and their properties**

	ALPHA α	BETA β	GAMMA γ
Nature of emission	${}^{4}_{2}\mathrm{He}$ (a helium nucleus)	${}^{0}_{-1}\mathrm{e}$ (an electron)	Very high frequency electromagnetic radiation
Charge	+2	−1	Zero
Penetrating ability — the ability to penetrate through a material before being stopped	Low — can only penetrate several centimetres of air and are stopped by a sheet of paper	Medium — can penetrate about one metre of air	High — can penetrate several centimetres of lead
Ionising ability — the ability to knock electrons from the atoms as it penetrates a material	High	Medium	Low



The varying ionising ability of these radiations is the basis of the operation of radiation detectors such as Geiger counters and scintillation tubes, as well as the discharge of electroscopes. Another way to distinguish between these radiations is by looking at the paths they follow through a magnetic or electric field. These are shown in figure 14.2. In each case the alpha and beta particles head off in different directions, due to their opposite charges, while the gamma radiation is unaffected. Also, because the beta particle is much lighter than an alpha particle, it follows a much tighter curve.

Figure 14.2 *The paths followed by different radiations in electric and magnetic fields*

PHYSICS FACT

Beta decay

The source of beta particles is a good example of scientific models improving with better technology. Beta particles are electrons like any other, except that they come from an atom's nucleus, not from an electron orbit. How is this possible?

The process of releasing a beta particle is known as beta decay, and it had traditionally been represented as in the following equation:

$$^{32}_{15}\mathrm{P} \rightarrow {}^{32}_{16}\mathrm{S} + {}^{0}_{-1}\mathrm{e} + \nu\ .$$

Here, ${}^{0}_{-1}\mathrm{e}$ is the electron or beta particle, while ν is an antineutrino. Note that the mass of the nucleus does not change significantly, although the identity of the atom does, due to its gaining a proton. After James Chadwick discovered the neutron in 1932, Enrico Fermi worked out a theory of beta decay that explained what was happening to the nucleons. As the following equation shows, a neutron turns into a proton, producing and ejecting an electron and an antineutrino in the process.

$$^{-1}_{0}\mathrm{n} \rightarrow {}^{1}_{1}\mathrm{p} + {}^{0}_{-1}\mathrm{e} + \bar{\nu}$$

After the 1964 proposal of the existence of quarks by Murray Gell-Mann and George Zweig, the search for quarks was pursued using particle accelerators. The last of the quarks to be found was the top quark, which was discovered in 1995 by physicists using the Fermilab accelerator in the USA. We now know that a neutron consists of two down and one up quark, while a proton consists of one down and two up quarks. The process of beta decay is now seen as a down quark turning into an up quark.

$$\mathrm{d} \rightarrow \mathrm{u} + {}^{0}_{-1}\mathrm{e} + \bar{\nu}$$

Turning a down quark into an up quark changes a neutron into a proton. This changes the identity of the atom without changing its mass significantly.

Binding energy

The mass of an atomic nucleus is less than the sum of the masses of its individual parts. Consider a helium nucleus (an alpha particle). It has a mass of 6.624×10^{-27} kg. It consists of two protons and two neutrons, and the sum of their masses is as follows:

$$\begin{aligned}
2\text{ protons} &= 2 \times 1.673 \times 10^{-27}\text{ kg} \\
2\text{ neutrons} &= 2 \times 1.675 \times 10^{-27}\text{ kg} \\
\text{sum of masses} &= 6.696 \times 10^{-27}\text{ kg.} \\
\text{Therefore, the mass difference} &= 6.696 \times 10^{-27} - 6.624 \times 10^{-27} \\
&= 7.2 \times 10^{-29}\text{ kg.}
\end{aligned}$$

This missing mass is called the 'mass defect'. It has an energy equivalent, as described by Einstein's equation, and this can be calculated as follows:

$$\begin{aligned} E &= mc^2 \\ &= 7.2 \times 10^{-29} \times (3 \times 10^8)^2 \\ &= 6.480 \times 10^{-12}\ \text{J}. \end{aligned}$$

This amount of energy is called the 'binding energy', because this is the amount of energy that must be acquired by a helium nucleus for it to split up into its separate protons and neutrons. The greater the binding energy per nucleon, the more stable a nucleus is. The atom with the most stable nucleus is element 26, iron (Fe), because it has the greatest binding energy per nucleon of all. A graph of the binding energy per nucleon is shown in figure 14.3.

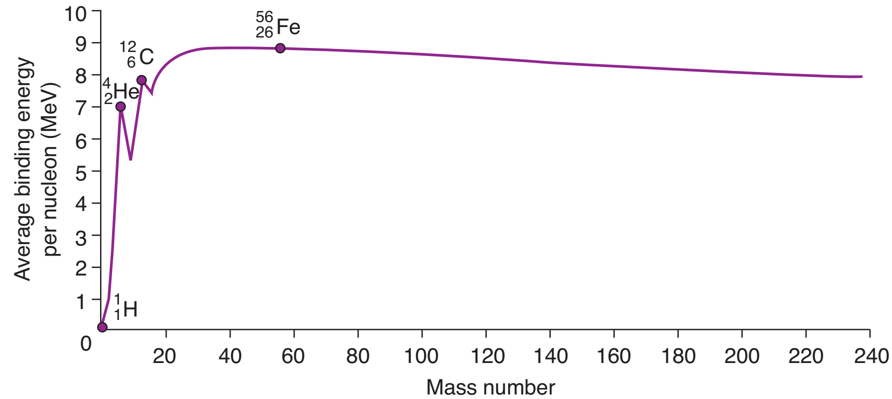


Figure 14.3 ***Binding energy per nucleon curve***

Any nuclear reaction that produces new elements with a greater binding energy per nucleon will result in mass being converted into energy and released. One way to do this is by splitting heavy nuclei, such as uranium, into smaller fragments. This is known as nuclear fission and is the process used in nuclear power stations and reactors such as the Australian Nuclear Science and Technology Organisation (ANSTO) Open Pool Australian Lightwater (OPAL) research reactor at Lucas Heights. Another way to get this energy is by joining small nuclei to form a slightly larger one. This is known as **nuclear fusion** and is the process that occurs inside a thermonuclear bomb blast, as well as inside stars.

Nuclear fusion is the joining of nuclei to form larger ones.

14.2 THE SUN

The Sun is a second-generation star. Evidence for this is in its composition — in addition to its 73.4% hydrogen and 25% helium, 1.6% of its mass is heavier elements. These are elements that were not produced in the big bang and can only have been produced in a large red-giant star, either during its lifetime or during its supernova. At 1.99×10^{30} kg the Sun has 329 000 times the mass of Earth; and with a diameter of 1 392 000 kilometres, it is 109 times the Earth's diameter.

Warning! You must never use a telescope to look directly at the Sun. To do so would place the retina of your eye at great risk. Solar observation is done using indirect methods. The simplest way for you to do this is to direct the image from a telescope onto a piece of paper.

Figure 14.1 shows an image of the Sun obtained using a telescope specifically designed for this job. Dark spots, known as sunspots, can be seen on the image. By observing the sunspots' movement across the face of the Sun, the period of rotation of the Sun has been measured. It was a surprise to discover that the Sun rotates faster at its equator, taking approximately 25 days to complete a rotation. Further from the equator the rotational period is longer, and at the poles it is estimated to be about 37 days. The average period of rotation of the Sun is usually taken to be 27.3 days.

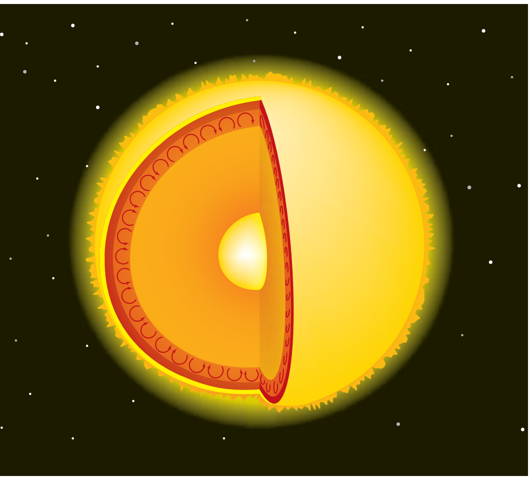


Figure 14.4 *The structure of the Sun*

The matter within the Sun is extremely hot and highly ionised plasma. The moving charges within the plasma create their own magnetic field, which interacts strongly with the Sun's magnetic field. A result of this is that the Sun's magnetic field lines become locked into the plasma, and flow with it. As the Sun rotates, the plasma at the equator moves past the slower rotating plasma at the poles and causes the field lines to twist and buckle. This can result in turbulent events, such as sunspots and solar flares, on the surface of the Sun.

Structure of the Sun

Close observation of the Sun has revealed shockwaves rippling across its surface. Analysis of the progress of these waves has revealed that the Sun has an inner structure made up of several layers. (A very similar technique has been used to deduce the internal structure of the Earth.) These layers, shown in figure 14.4, are:

- *the core.* The innermost layer and site of energy production by the fusion of hydrogen to helium.
- *the radiative zone.* Electromagnetic energy is transmitted slowly through this layer by successive absorption and re-radiation.
- *the interface zone.* This thin layer seems to generate the Sun's magnetic field.
- *the convection zone.* Energy is transmitted through to the surface by convection currents.

The **photosphere** is the visible surface of the Sun.

The visible surface of the Sun is called the **photosphere**; however, the surface is not like the Earth's. The photosphere is a gaseous layer several hundred kilometres thick with a density of about 10^{17} atoms per cubic centimetre. That is about one ten-thousandth of the density of the atmosphere at the Earth's surface. We see this layer as the 'surface' because, at a temperature of 6000 K, most of the electromagnetic radiation it produces is in the visible spectrum.

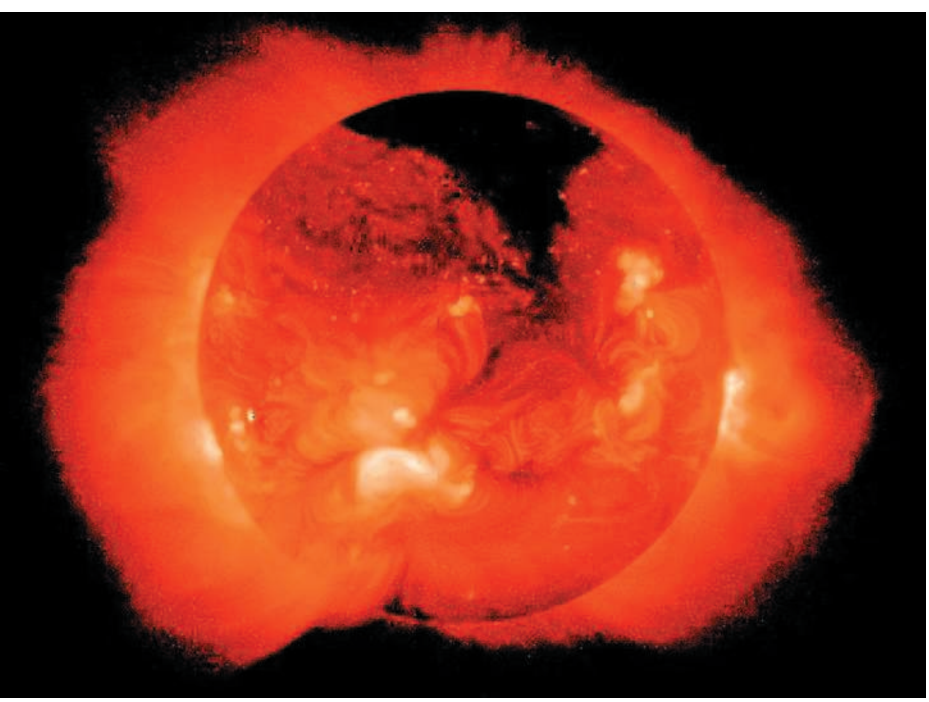


Figure 14.5 *The Sun's corona, an X-ray photograph produced by the Japanese* **Yohkoh** *satellite*

Above the photosphere is the atmosphere of the Sun, which is normally invisible but can be seen during a total solar eclipse. The Sun's atmosphere is also made up of several layers:

- *the chromosphere* — immediately above the photosphere
- *the transition region* — just a few hundred kilometres thick
- *the corona.* This is the outer atmosphere of the Sun, and extends many millions of kilometres out into space. It does not end abruptly, but streams out into space, forming the solar wind. Its temperature is 2 000 000 K at an altitude of 75 000 kilometres above the photosphere, although its density is extremely low. The corona has been found to be a strong X-ray emitter. This means that it can be photographed using these wavelengths. This is done regularly by a Japanese satellite called *Yohkoh*, which produced the image shown in figure 14.5.

Emissions from the Sun

The Sun produces two types of emissions. The first is electromagnetic radiation and the other is the solar wind.

Electromagnetic radiation

The electromagnetic spectrum has already been covered earlier in this text. The Sun produces a range of electromagnetic wavelengths, from the short-wavelength gamma rays to the long radio waves. However, the visible spectrum is the most intense wavelength range produced. The wavelength of maximum intensity is about 460 nanometres. Electromagnetic radiation travels at a speed of 3×10^8 m s^{-1}, which means that it takes a little over eight minutes to reach the Earth from the Sun.

PHYSICS FACT

The Sun as a black body

The radiation from the Sun in the visible and infra-red wavelengths closely matches the theoretical radiation curve of a black body with a temperature of approximately 6000 K. This portion of the solar spectrum is produced by the photosphere, which has that temperature. At shorter and longer wavelengths the solar spectrum is different, corresponding instead to the spectrum of a black body at a temperature of approximately 1 000 000 K. This portion of the solar spectrum is produced by the hot corona and by solar flares. At ultraviolet and low-intensity X-ray wavelengths, the solar spectrum does not match either of these two theoretical black-body radiation curves.

The solar wind and interplanetary magnetic field

The solar wind is an outflow of low-density plasma from the corona of the Sun.

The difference in pressure between the corona and interplanetary space (that is, the space between the Sun and the rest of the solar system) causes an outflow of material from the corona. The low-density, hot plasma of the corona streams almost directly outward from the Sun, eventually reaching each of the planets. This **solar wind** has a speed of 400–500 kilometres per second, with occasional bursts of material travelling at up to 1000 kilometres per second. At these speeds, the wind takes three to four days to reach the Earth. Near the Earth, the solar wind has a density of approximately five protons and five electrons per cubic centimetre, with some helium and heavier ions (the solar wind is approximately 8% He). The velocity of the particles gives them a temperature of between 10 000 K and 100 000 K.

As happens in the photosphere, a magnetic coupling occurs between the magnetic field belonging to the plasma of the solar wind (recall that moving charges generate a magnetic field) and the Sun's magnetic field, with the result that the Sun's magnetic field lines become 'frozen' (or locked) into the plasma. Wherever the plasma flows, it carries the magnetic field lines with it. Therefore, the outflowing solar wind carries the Sun's magnetic field out with it to form the interplanetary magnetic field. The magnetic field lines remain anchored within the surface of the Sun and point out into interplanetary space. However, as the Sun rotates, it winds the field lines up into a giant spiral as shown in figure 14.6. At the Earth, the field lines are at approximately 45° to the direction of the Sun, while further out in the solar system, the field lines are almost perpendicular to the direction of the Sun.

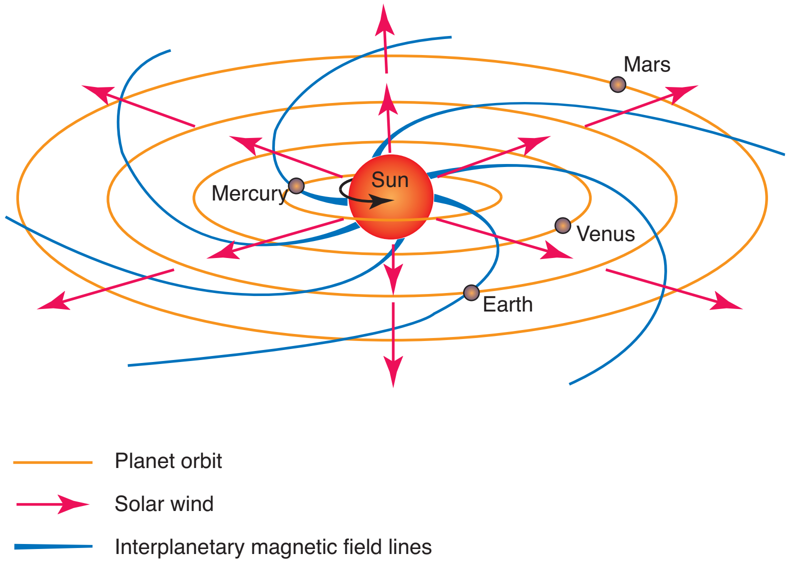


***Figure 14.6** The field lines of the interplanetary magnetic field emanate perpendicularly from the surface of the Sun but become wound up into a giant spiral.*

The solar wind is the subject of observations by many different solar probes and observatories. The *Advanced Cosmic Explorer (ACE)* has been placed at the L1 point. The L1 point is a stable position about one-hundredth of the distance from the Earth to the Sun that allows a spacecraft to orbit the Sun together with the Earth. *ACE* continually samples the solar wind and can report on approaching gusts before they reach the Earth. The space probe *Ulysses* was placed into a solar orbit perpendicular to the plane of the solar system (the ecliptic), allowing it to measure the solar wind over the Sun's poles. The results of this survey concluded that the solar wind principally originates in coronal holes, which can usually (though not exclusively) be found over the solar poles. Holes in the Sun's corona can be seen in figure 14.5 on page 291.

Features of the Sun

An examination of the surface and atmosphere of the Sun reveals a number of discernable features. The photosphere can be examined in visible light; however, the atmosphere is not normally visible and can only be seen in a few ways. These include observation during solar eclipses, close observation at the edge of the image of the Sun (called the limb), and observation made in X-ray wavelengths. The types of feature that can be seen are:

- granulation. Seen in figure 14.7, the photosphere appears to be covered in granules, approximately 500–1000 kilometres across. Each one of these granules is the top of a convection current lying beneath, within the convection zone. They have an average lifetime of just eight minutes as the currents shift and change.

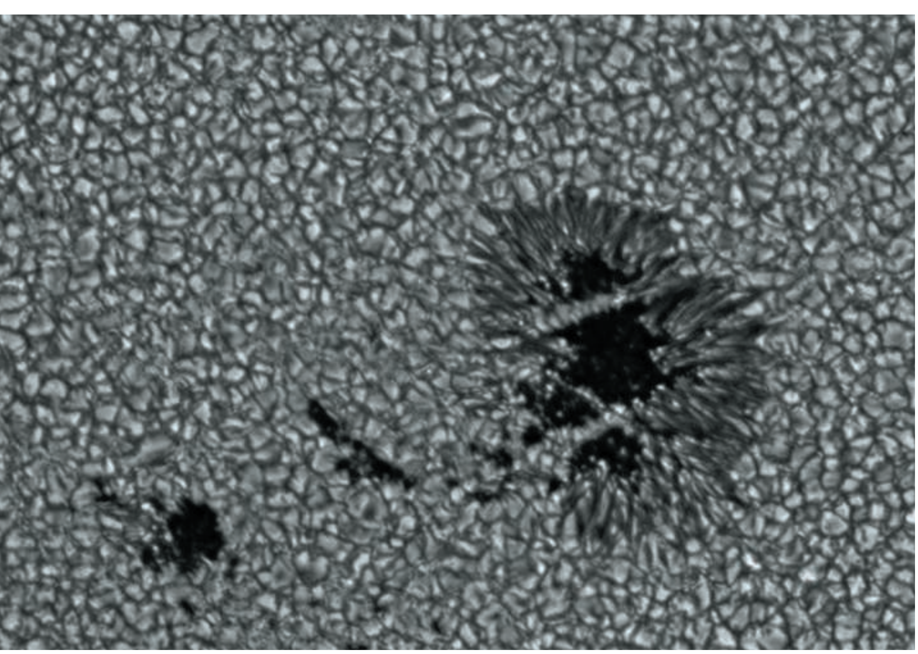


***Figure 14.7** A photograph of the surface of the Sun showing granulation as well as sunspots*

- spicules. These short-lived jets of gas, shown in figure 14.8, rise several thousand kilometres from the surface and are seen around the edges of sunspots and granules.

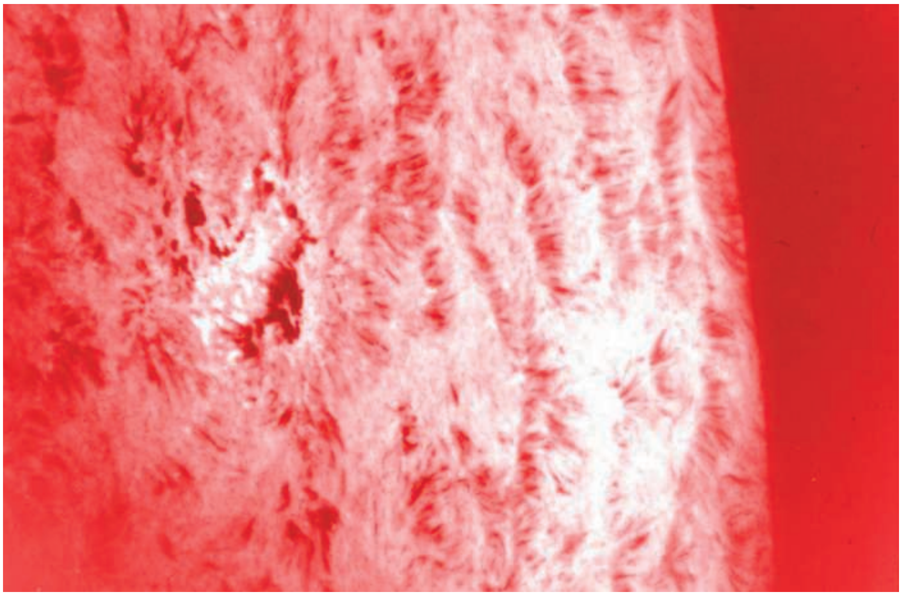


***Figure 14.8** Spicules, short-lived jets of gas seen around the edges of sunspots and granules*

A sunspot is a dark spot on the surface of the Sun, representing a region of intense magnetic activity and lower temperature.

- **sunspots**. These are dark spots seen on the surface of the Sun, varying in size between several hundred to several thousand kilometres in diameter. Examples are shown in figure 14.1 and closer in figure 14.7. Sunspots have been identified since Galileo first observed the Sun using a telescope (projecting the image onto paper). They appear dark because they are about 1500 K cooler than their surroundings. The spots also represent regions of intense magnetic activity, containing magnetic fields up to 0.4 tesla (compare this to the Sun's usual magnetic intensity of just 1×10^{-4} tesla, or the Earth's 3×10^{-5} tesla). It is thought that sunspots are locations of disturbances in the magnetic field lines within the surface of the Sun, where they have become sufficiently buckled to loop out and then back into the surface. The intense field activity within a sunspot prevents the convection of heat to the surface, thereby reducing its temperature. Sunspots usually occur in pairs or groups, lasting for several days or weeks.
- flares. Shown in figure 14.9, solar flares are sudden explosive outbursts of radiation and matter that occur near sunspots and last about an hour. The radiation bursts are within the radio waves to X-ray range, while the outburst of matter into the solar wind, called a coronal mass ejection, can have an impact on the Earth. Solar flares are the result of magnetic field lines becoming so wound up and twisted that they snap and rejoin, with an associated sudden release of energy. As a result, solar flares are more common around the more complex sunspot groups.

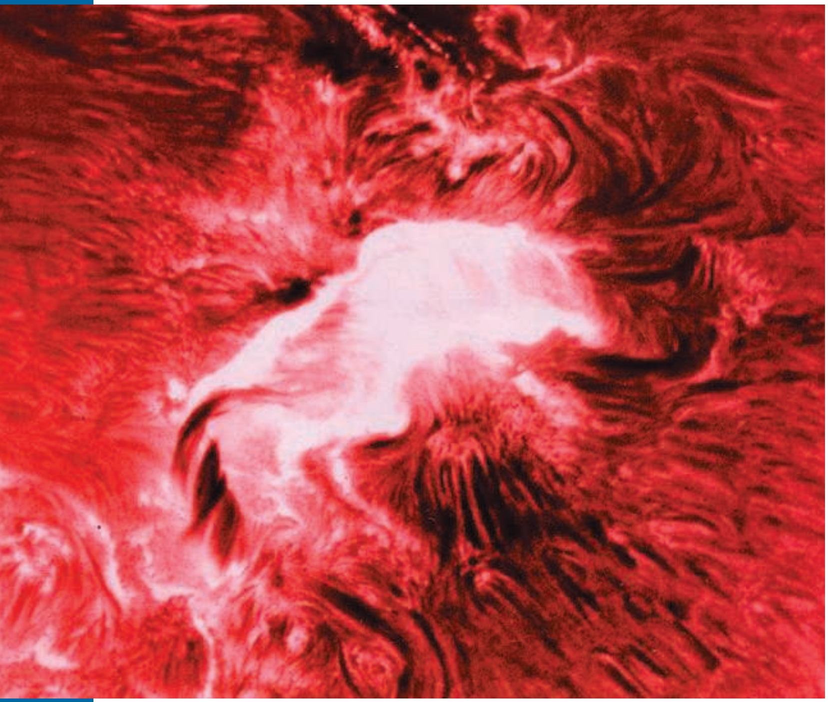


***Figure 14.9** A close-up view of a solar flare*

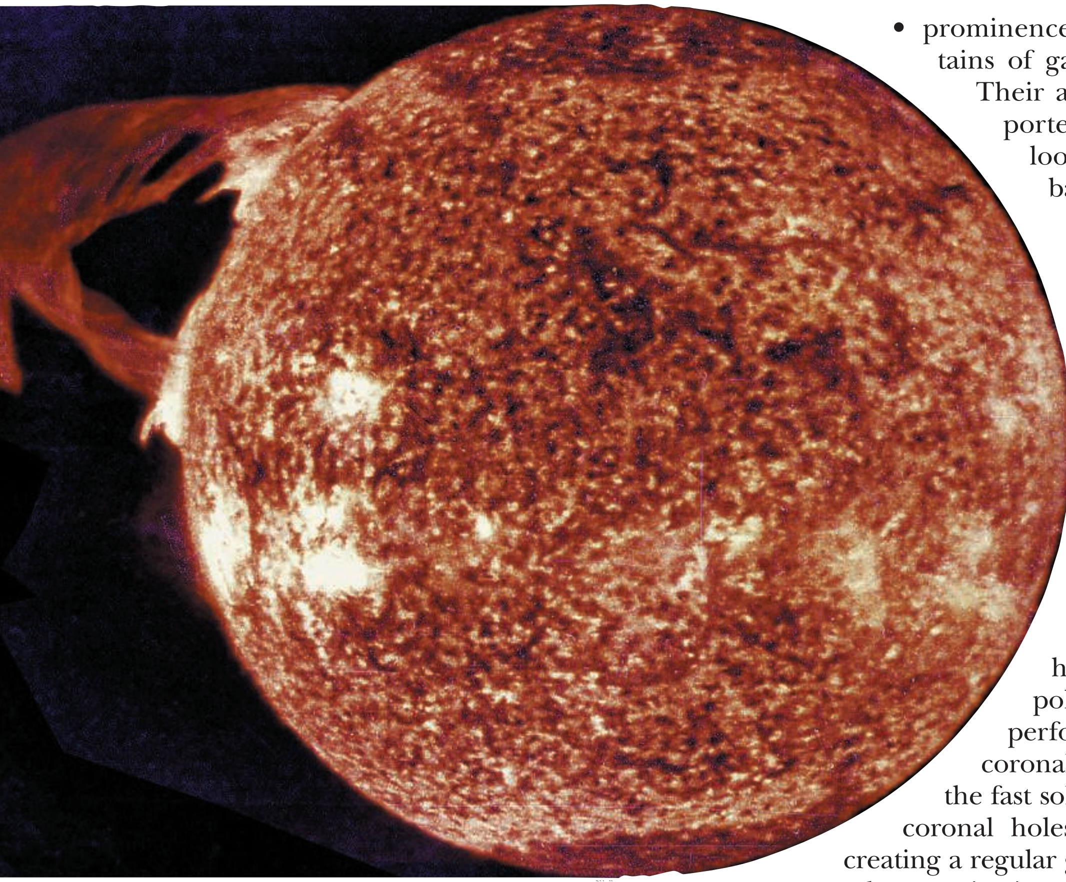


Figure 14.10 *One of the largest solar prominences ever recorded on film, rising 600 000 km above the surface of the Sun*

- prominences. These are large, looping curtains of gas that can last for several days. Their arched shape follows, and is supported by, magnetic field lines that loop out of the surface and then back in. Prominences can also cause material to be thrown out into the corona, to become fluctuations in the solar wind. Figure 14.10 shows one of the largest prominences ever recorded on film.
- coronal holes. The corona does not wrap completely around the Sun, but usually contains holes, particularly over the solar poles. Looking back to the X-ray photograph of the corona shown in figure 14.5, enormous coronal holes can clearly be seen over the poles. The survey of the solar wind performed by *Ulysses* has found that coronal holes are the primary source of the fast solar wind. As the Sun rotates, large coronal holes reappear every 27 days or so, creating a regular gusting in the solar wind.
- coronal mass ejections. The existence of coronal mass ejections (CME) has only been known since the early 1970s. CMEs are large, magnetically confined bubbles of plasma, containing between one and ten million million kilograms of coronal material. They expand and accelerate away from the Sun, reaching speeds up to 1000 kilometres per second. CMEs are released by solar flares and prominences, as well as some other solar events. At their slowest rate of release, one CME per week is recorded, while at peak production two or three per day can be found.

The onrush of large volumes of fast moving plasma can have a severe impact on the Earth. For example, the 1997 failure of AT&T's Telstar satellite coincided with the arrival of a CME that had erupted from the Sun two days earlier.

The solar cycle

It has already been mentioned that the Sun experiences periods of turbulent activity and periods of relative quiet. It is a cyclical pattern, which repeats every 11 years on average. The period of peak activity is called the solar maximum, while the period of least activity is referred to as the solar minimum. As activity increases after a minimum, the frequency of sunspots, flares, prominences and coronal mass ejections all increase, along with their associated disruptions to the solar wind. The easiest of these features to observe are sunspots, and so the **sunspot cycle** is closely monitored by several agencies worldwide in order to track the more general **solar cycle**. This practice began at the Zurich Observatory in 1749. In Australia, this function is currently performed by IPS Radio and Space Services, who issue regular bulletins and alerts, and provide information through their website.

The sunspot cycle is a cyclical pattern of increasing and decreasing numbers of sunspots.

The solar cycle is an 11-year cyclical pattern of increasing and decreasing frequency of sunspots, flares, prominences and coronal mass ejections.

From these observations, it has been found that:

- at solar maximum there is usually over 100 sunspots appearing simultaneously on the surface of the Sun, while at solar minimum there may be none at all
- although the average cycle length is 11 years, the cycle length varies between 7 and 13 years
- an average cycle is not symmetrical, usually taking 4 years from minimum to maximum, and 7 years to quieten down to the next minimum.

Because the cycle length varies slightly it can be difficult to pick the start of a new cycle, which is defined to begin at a minimum. Two methods have been devised to track each cycle. The first is to count the number of sunspots appearing daily and then to graph this data, as shown in figure 14.11. Over a period of time the cycle becomes apparent. Figure 4.12 shows a graph of the accumulated data since sunspot numbers have been monitored. Note that the solar cycle occurring at the start of observations in 1749 is referred to as 'sunspot cycle zero'. By counting along the graph it is easily seen that at around the middle of the year 2000 we experienced the maximum of sunspot cycle 23.

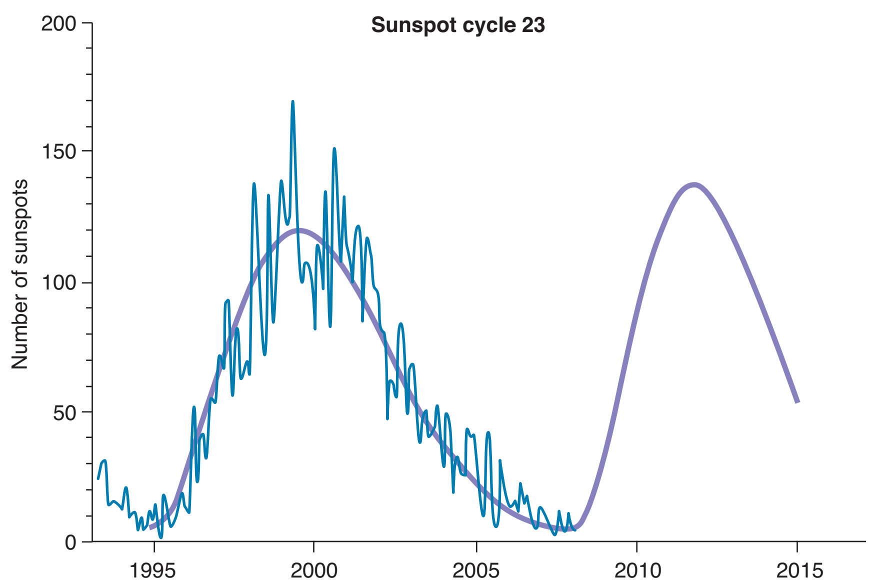


Figure 14.11 *By plotting the sunspot numbers observed each day the progress of the sunspot cycle can be followed. The plot shows sunspot cycle 23 finishing at the end of 2007 and the smooth curve shows the prediction for cycle 24.*

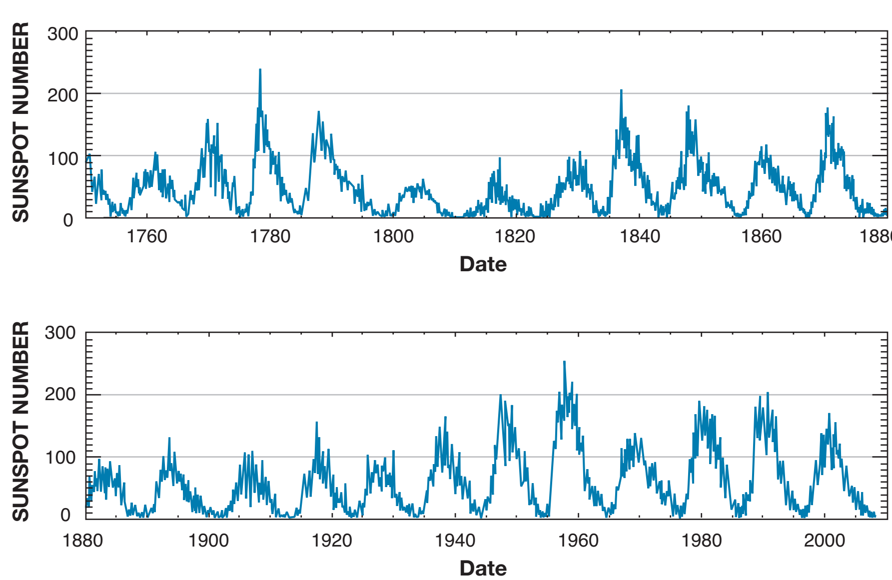


Figure 14.12 *A graph of sunspot numbers since observations began in 1749. The repetitive pattern of the sunspot cycle is clearly visible.*

The second method to track the progress of a sunspot cycle is to note the latitude at which sunspots are occurring. It has been found that, early in a cycle, sunspots appear at higher latitudes, typically about 40° (although they have been seen as high as 60°). As the cycle progresses, the sunspots appear closer to the solar equator. Near the end of the cycle they are located close to the equator, at a latitude of about 5°. By plotting the latitudes of sunspots over time, a pattern is revealed that has become known as a butterfly diagram. A butterfly diagram is shown in figure 14.13.

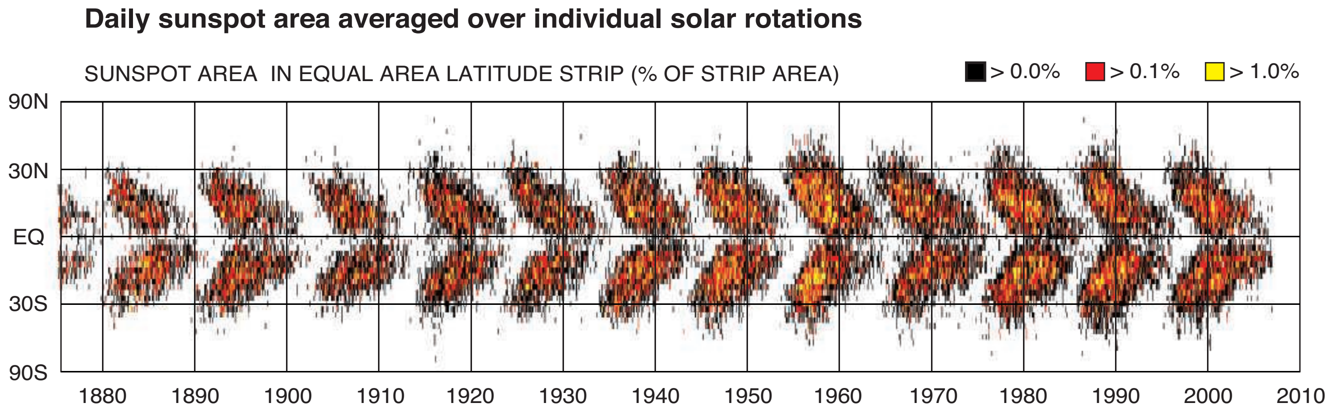


Figure 14.13 *A butterfly diagram, showing the way that sunspots appear at high latitudes early in a cycle and progress towards the equator. This pattern is helpful in tracking the progress of a cycle.*

14.3 THE SUN–EARTH CONNECTION

Recall that there are two types of emission reaching the Earth from the Sun — electromagnetic radiation and the solar wind. Each of these have an impact on the Earth, and in each case the Earth's atmosphere shelters us from excessive exposure to these emissions.

The electromagnetic connection

The Sun emits a wide range of electromagnetic wavelengths, from gamma rays to radio waves, and all of these impinge on the outer atmosphere of the Earth. However, not all of these different wavelengths manage to penetrate the atmosphere and reach the surface of the Earth. Different wavelengths get absorbed by different atmospheric molecules, located at different altitudes. Gamma rays and X-rays are absorbed by oxygen and nitrogen at an altitude of about 100 km, while ultraviolet radiation is absorbed by ozone at about 100 km. Infra-red wavelengths are absorbed by water vapour and carbon dioxide at 20 km altitude. This information is shown in figure 14.14. Note that there are two portions of the electromagnetic spectrum that are not absorbed — the visible and near infra-red, as well as microwaves and radio waves. As the atmosphere is transparent to these two broad bands of wavelengths, they are referred to as selective transmission windows.

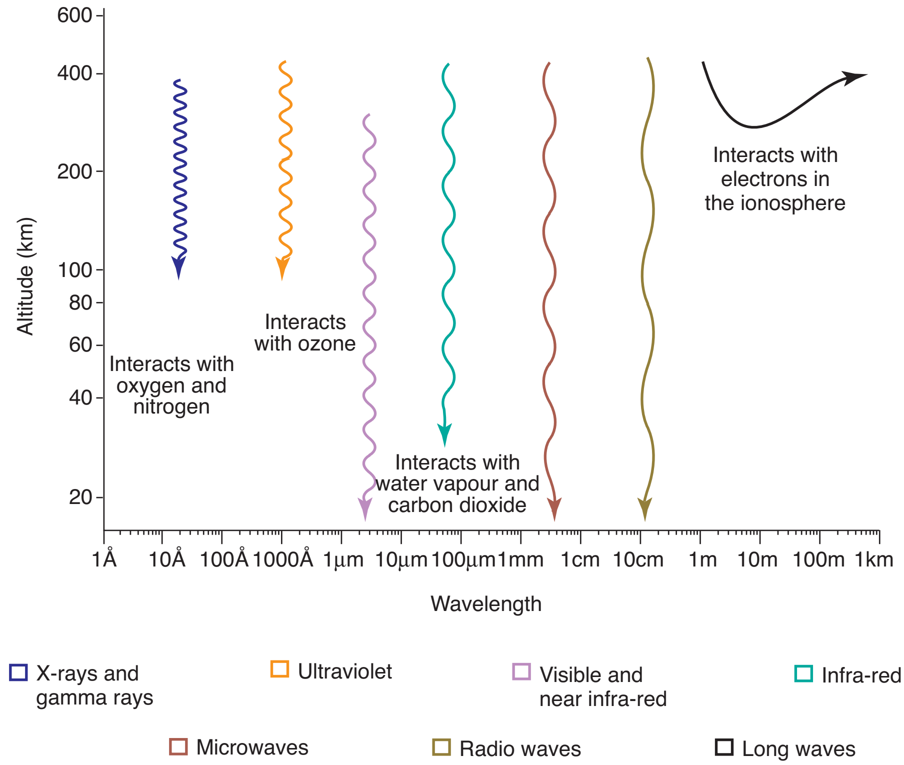


Figure 14.14 ***Absorption and transmission of different portions of the electromagnetic spectrum***

The absorption of higher frequency gamma rays, X-rays and ultraviolet rays affords life on Earth a level of protection from these potentially harmful radiations. In addition, transmission of visible light and near infra-red allows some solar energy to be delivered to animal and plant life on or near the Earth's surface. The absorption of portions of the infra-red radiation by water molecules also provides a warming mechanism for the atmosphere.

The solar wind connection

The solar wind takes approximately three to four days to reach the Earth, although high-speed gusts can reach us in two days. Recall that the solar wind is a flow of highly ionised charges carrying the interplanetary magnetic field with it. The interplanetary magnetic field couples with the Earth's magnetic field, allowing the solar wind to interact with the Earth's magnetic field. Most of the solar wind flows around and past the Earth, distorting the Earth's magnetic field (which would otherwise have the simple shape of a bar magnet's field) to create a long tail. The region containing a planet's magnetic field, distorted in this way, is known as a **magnetosphere**. It is shown in figure 14.15. Mercury, Jupiter, Saturn, Uranus and Neptune also possess magnetospheres.

A magnetosphere is the region surrounding a planet that contains its distorted magnetic field.

The ions of the solar wind can enter the magnetosphere in three ways:

- through the cusps, holes in the magnetosphere, over the north and south poles
- through the magnetotail and then back up towards the Earth
- through some leakage in the magnetopause.

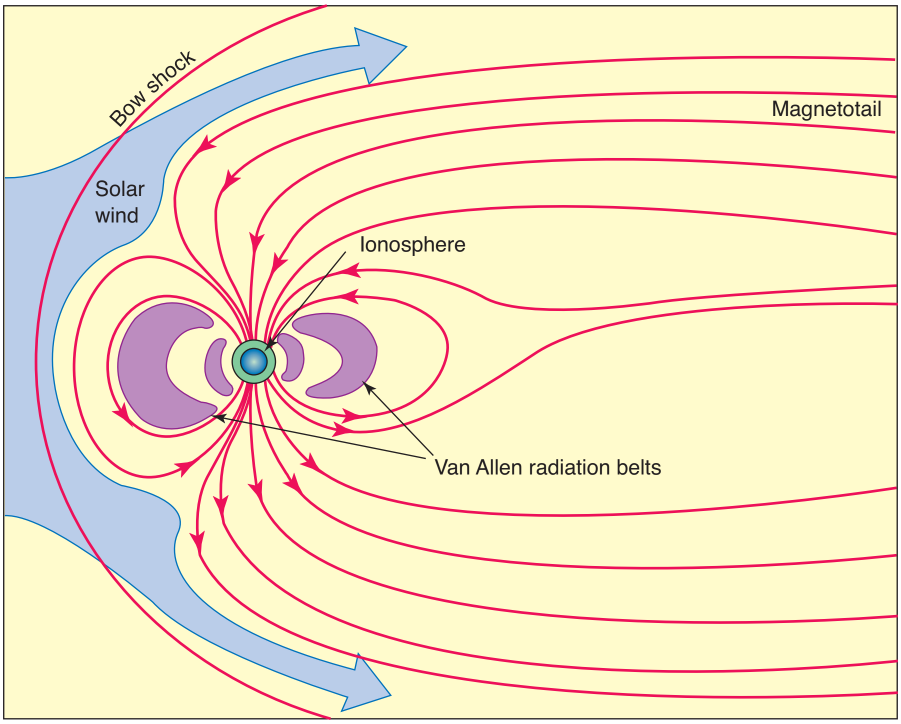


Figure 14.15 *The magnetosphere of the Earth. Note the flow of solar wind against and around the magnetosphere.*

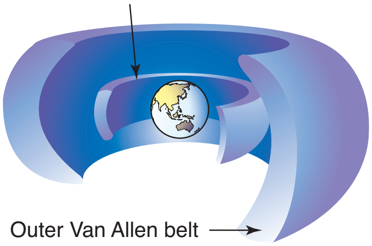


Figure 14.16 *The inner and outer Van Allen belts*

The Van Allen belts are two doughnut-shaped zones of radiation (ions) that wrap around the Earth.

However, once inside, the ions are captured by the Earth's magnetic field lines. They spiral from one pole to the other and are then bounced back, this round trip taking just one-tenth of a second. The captured ions accumulate into the outer of two doughnut shaped zones that wrap around the Earth. (The inner zone, possessing higher energy particles, is much more stable and is the result of a slow accumulation of charges derived from cosmic rays, which are high-energy particles that bombard us from all directions, not from the Sun.) Shown in figure 14.16, these zones are known as the **Van Allen belts**. They were discovered by James Van Allen in 1958 at the very beginning of the USA's space-rocket efforts.

The charged particles within the outer Van Allen belt have been found to follow a third motion, in addition to their field-line spiral and pole-to-pole bounce. The charges also skip from one field line to an adjacent one, drifting sideways as they bounce from pole to pole. Positive ions have been found to move in a westward direction, while electrons move eastward. This amounts to an electrical current, called the ring current, running around the Earth. Electric currents produce their own magnetic fields, and the magnetic field of the ring current interferes with the Earth's magnetic field. During times of turbulent solar winds, the ring current fluctuates, and so too does its influence on the Earth's magnet field. This contributes to the effects of a **geomagnetic storm**, during which the Earth's magnetic field experiences unusual distortions and fluctuations in strength.

A geomagnetic storm is a period during which the Earth's magnetic field experiences unusual distortions and fluctuations in strength.

Space weather and its effects

The magnetosphere offers some protection for our planet from direct exposure to the solar wind; however, it is still vulnerable to major disruptions in the speed, density and temperature of the solar wind. These are the conditions referred to as space weather. Turbulent solar winds buffet and irregularly distort the magnetosphere, and at the same time a greater than usual number of charged particles find their way into the magnetosphere to irregularly charge the Van Allen belts and ionosphere. The action of stormy space weather on the Earth and its magnetosphere produces a range of side effects, including:

- abnormal heating of the atmosphere, causing it to expand slightly and increase drag on low satellites (which can result in premature decay of their orbits, causing them to re-enter the Earth's atmosphere and, most probably, burn up)
- higher than usual risk of radiation exposure for astronauts, satellites and spacecraft
- electrical failure of communications satellites
- strong geomagnetic storms that can produce auroral displays, but also send huge current spikes through power grids, resulting in their failure
- fluctuations in the Van Allen belts and the ionosphere resulting in disruptions to normal communications.

Part of the function of the *ACE* space probe has been to provide an early warning system (about one hour) of impending bad space weather. The probe uses four instruments to measure the solar-wind particle flow (speed, density and temperature) as well as the interplanetary magnetic field strength and direction. Reports of current space weather conditions as measured by *ACE*, such as the real-time display shown in figure 14.18 and the 'browse plot' in figure 14.17, are readily available from websites such as 'Space Weather Now' run by the US Space Weather Prediction Center, or www.spaceweather.com.

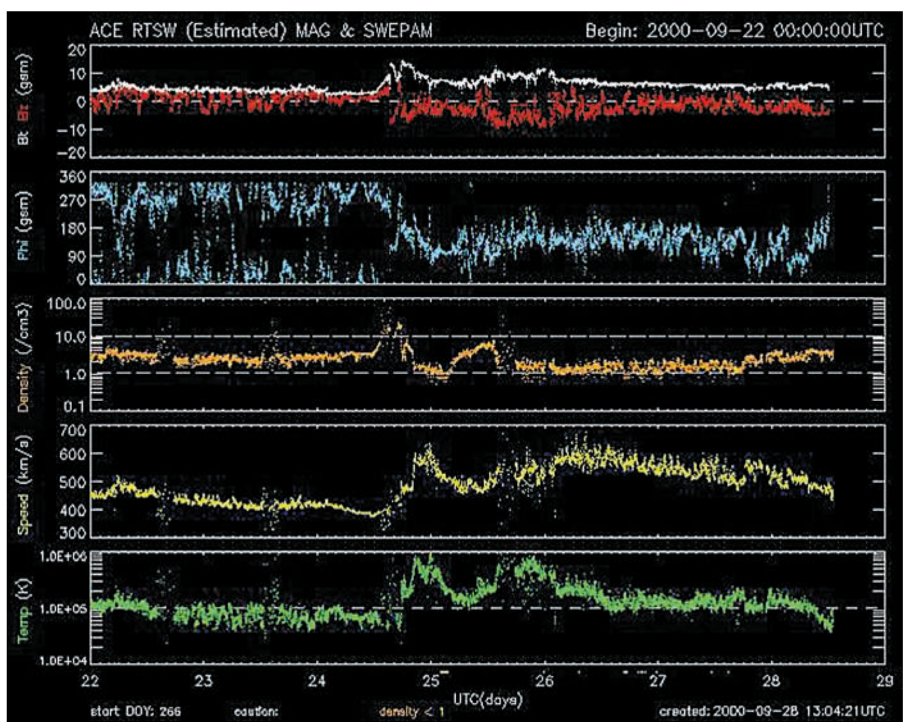


Figure 14.17 ***A space weather report from the* ACE *space probe, covering several days during the Sydney Olympics, 2000. The two upper graphs indicate interplanetary magnetic field fluctuations, while the lower three graphs indicate fluctuations in the solar-wind density, speed and temperature. Note that an event occurred on 25 September, the date of Cathy Freeman's 400 metre gold-medal run.***

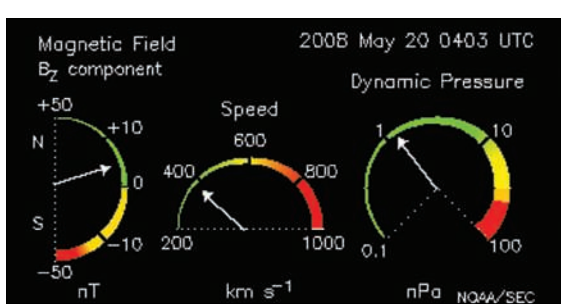


Figure 14.18 ***A real-time space weather report***

PHYSICS FACT

Adverse effects of the solar maximum

The turbulent solar winds during a solar maximum produce geomagnetic storms within the magnetosphere that can have an impact on human activity in a variety of ways. The following are specific examples.

In January 1997, a coronal mass ejection headed straight at the Earth. As it struck the Earth's magnetosphere, on 11 January, it produced a geomagnetic storm. A US$200 million communications satellite belonging to AT&T abruptly ceased operating. It had been working smoothly for two years and was expected to last much longer. Review teams eventually concluded that the failure of a simple power-supply diode was to blame.

In mid-March 1989, a strong geomagnetic storm raged around the globe. At 2.44 am on 13 March, induced electrical currents created by the geomagnetic storm caused one of the transformers in the power transmission grid of HydroQuebec in Canada to fail. This event precipitated the collapse of the entire power grid. The process was quick, taking just 90 seconds to plunge 6 million people into a blackout that lasted 9 hours.

On 14 May 1973, the US launched a space station called *Skylab* into a 435-km-high orbit of the Earth, intending it to stay there for at least ten years. In 1979, a strong solar maximum occurred. Because of the additional heating and expansion of the atmosphere, *Skylab* experienced unanticipated atmospheric drag that degraded its orbit. NASA fought hard to keep its space station aloft, but *Skylab* soon made an uncontrolled re-entry. It was much too big to burn up during re-entry, and crashed into the Indian Ocean, with some of the debris landing in Western Australia. Luckily, it avoided any population centres.

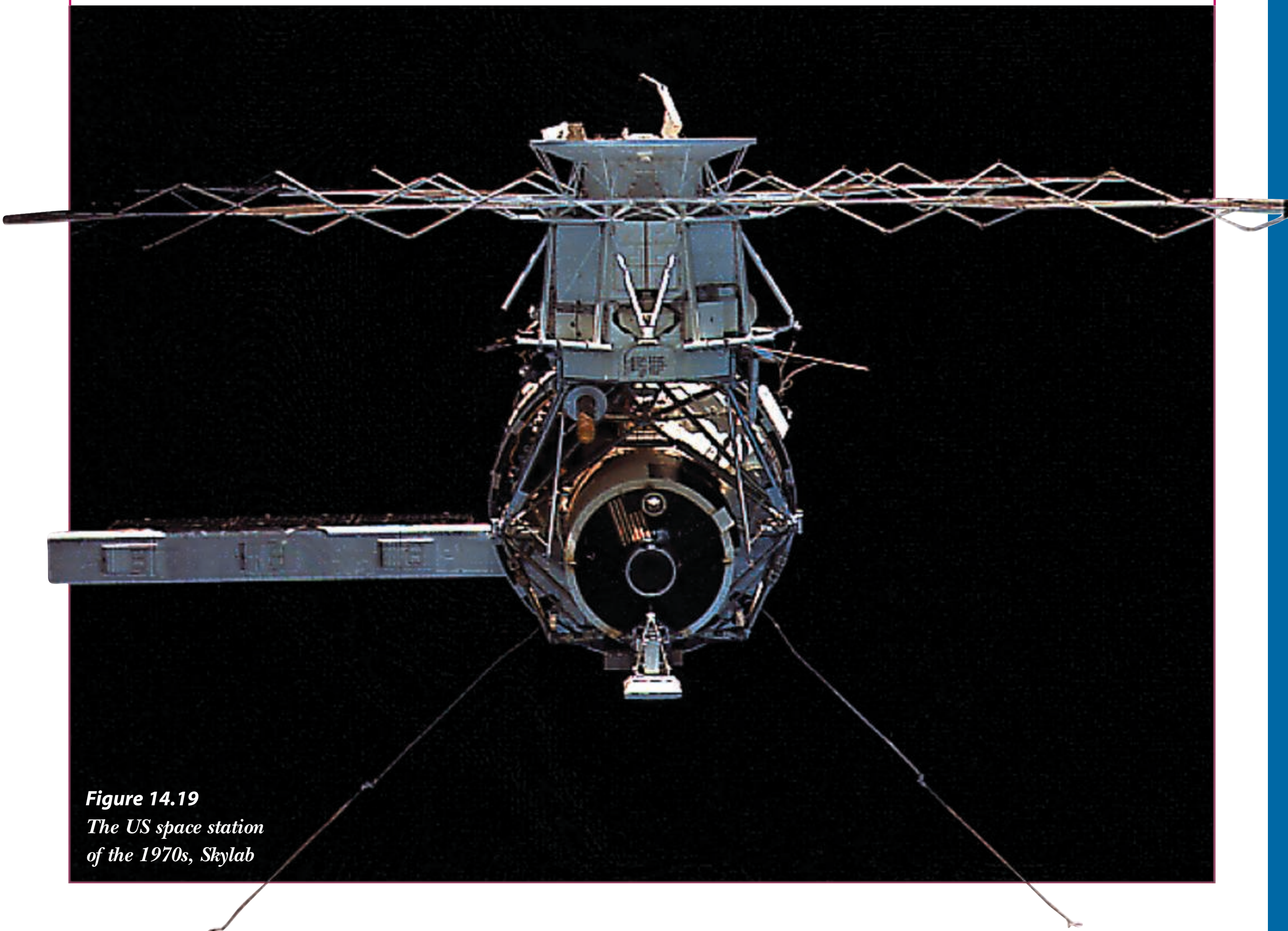


Figure 14.19
The US space station of the 1970s, Skylab

PHYSICS IN FOCUS

The auroras

The charges within the outer Van Allen belt can descend down to approximately 100 km near the poles during their north–south motion. As they descend they collide with tenuous upper-atmospheric molecules, mostly oxygen with some nitrogen. These collisions give energy to the particles, which then re-emit the energy at their own peculiar frequencies, producing light. These lights are known as auroras — near the north pole they are called aurora borealis, while near the south pole they are called aurora australis. Figure 14.20 (below) shows an example of an aurora. The enchanting lights build and fade and move around, taking on a variety of colours. Green light is produced by collisions with oxygen atoms at an altitude of about 100 km; blue light is produced by ionised nitrogen; purplish-red by neutral nitrogen; and red light is produced by sparsely distributed oxygen atoms at a very high 300-km altitude. Because the outer Van Allen belt surrounds the Earth, the auroras occur in a ring around the polar caps. This is called the auroral oval, and is shown in figure 14.21 (right). During magnetic storms the auroral oval expands and auroras can be seen further from the poles than usual.

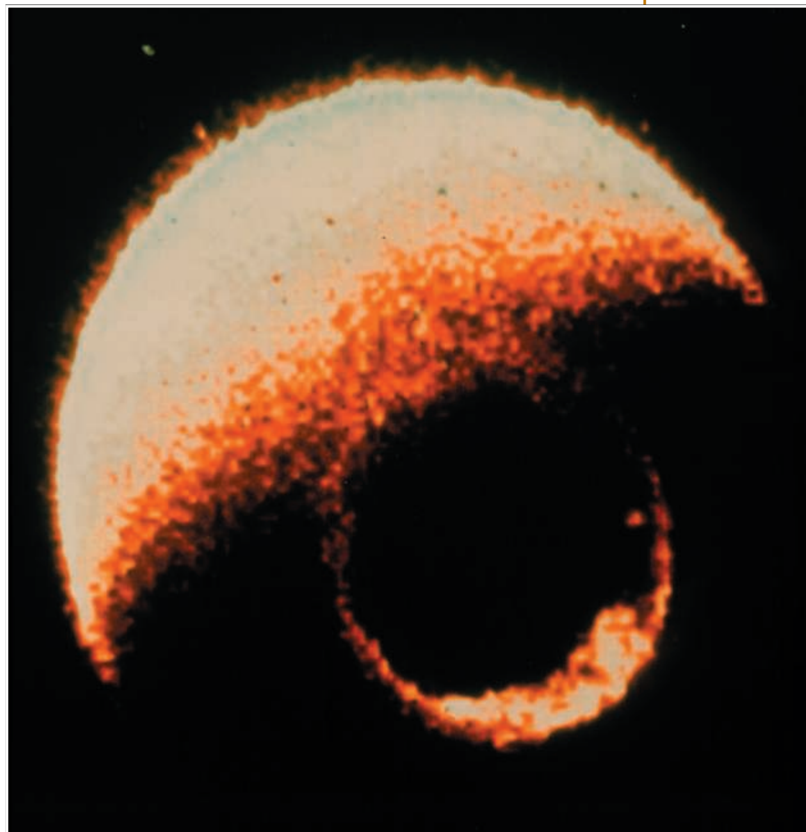


Figure 14.21 *The auroral oval, as seen from space*

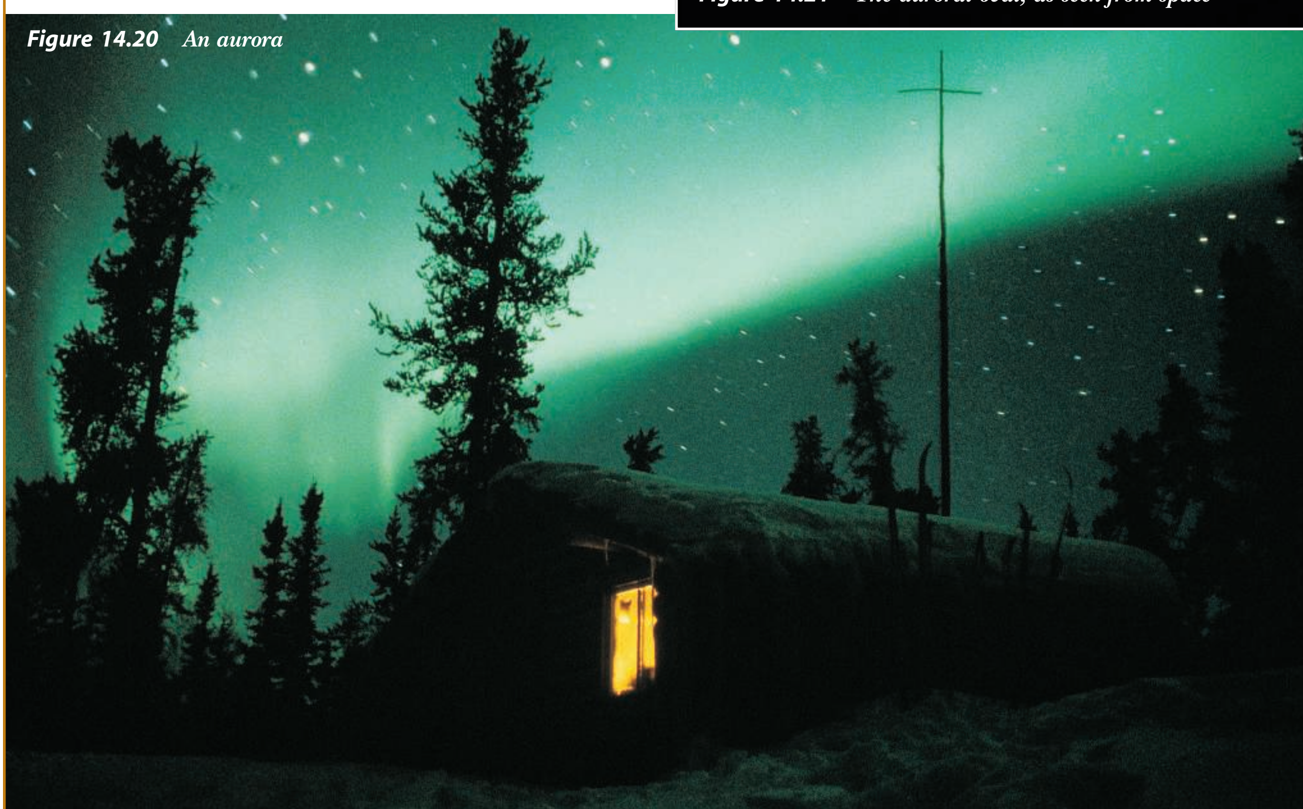


Figure 14.20 *An aurora*

SUMMARY

- There are three types of nuclear radiation: alpha α (helium nucleus), beta β (electron) and gamma γ (high-frequency electromagnetic radiation).
- These vary in ionising and penetrating power, as well as in the effect of electric and magnetic fields upon them.
- Any nuclear reaction that results in an increase in binding energy per nucleon will release energy. This includes fission (splitting) of large nuclei and fusion (joining) of small nuclei.
- The matter within the Sun is extremely hot and highly ionised, forming a state of matter known as a plasma.
- Emissions from the Sun include electromagnetic radiation, as well as the solar wind with the interplanetary magnetic field.
- The Sun displays a number of features, including granulation, spicules, sunspots, flares, prominences, coronal holes and coronal mass ejections.
- Sunspots are regions of lower temperatures and strong magnetic activity.
- Many of the features of the Sun increase and decrease in frequency according to the 11-year solar cycle. The corresponding periodic variation in the number of sunspots is called the sunspot cycle.
- Electromagnetic radiation from the Sun is selectively absorbed by molecules within the Earth's atmosphere.
- The solar wind and interplanetary magnetic field link to, and distort, the Earth's magnetic field (called the magnetosphere). The magnetosphere deflects the solar wind around itself, but also manages to capture some of the ions of the solar wind.
- Solar wind ions captured by the Earth's magnetic field form the outer Van Allen belt. High-energy ions captured as a result of cosmic rays form the inner Van Allen belt. Currents within these belts produce magnetic effects of their own.
- During a solar maximum, the emissions from the Sun can fluctuate erratically. These fluctuations have an impact on the Earth, producing effects such as geomagnetic storms, possible satellite failure, possible power-grid failure, communication difficulties and unusual auroras.

QUESTIONS

The Sun

1. Complete the following table comparing various attributes of nuclear radiations.

	ALPHA α	BETA β	GAMMA γ
Nature of emission			
Charge			
Penetrating ability			
Ionising ability			

2. There are similarities between the behaviour of nuclear radiations in electric and magnetic fields.
   (a) Describe the effect of these fields on the path of gamma rays.
   (b) In what two ways do the paths of alpha and beta particles differ in these fields? Provide a reason for each difference.
3. (a) What is binding energy and how does it give a nucleus stability?
   (b) What types of nuclear reactions are exothermic — that is, they release energy?
   (c) Which of these types of reaction occurs inside a star?
4. (a) Briefly describe the nature of the state of matter known as plasma.
   (b) Plasma has the ability to hold magnetic field lines. Outline how this property leads to turbulent events on the surface of the Sun.
5. Describe the nature and location of the source of the Sun's energy.
6. All photographs of the Sun show a surface, and yet it has no surface like the Earth. Explain this apparent contradiction.
7. Identify the three layers of the Sun's atmosphere and state the difference between each layer.
8. Identify the two types of emission that reach the Earth from the Sun.

9. (a) Describe the nature of the solar wind.
   (b) Describe the relationship between the solar wind and the Sun's corona.
   (c) Describe the relationship between the solar wind and the Sun's extended magnetic field.
10. (a) List the various features that can be seen on the surface of the Sun. Include a brief description of each.
    (b) Write a more detailed description of sunspots, referring to temperature and magnetic activity.
11. (a) What is the sunspot cycle? Describe the pattern of this cycle.
    (b) Outline the link between the sunspot cycle and the solar cycle.
    (c) List other events linked to the solar cycle.

The Sun–Earth connection

12. Describe the protection provided by the Earth's atmosphere from the range of electromagnetic radiation reaching us from the Sun.
13. Sketch the shape of the Earth's magnetic field, distorted by the solar wind and interplanetary magnetic field. Include arrows to show the places where solar-wind particles can enter the magnetosphere.
14. (a) Outline the nature of a geomagnetic storm.
    (b) Explain how the motion of the charges within the Van Allen belts contributes to a geomagnetic storm.
15. (a) What is space weather?
    (b) Describe how the magnetosphere, the Van Allen belts and the ionosphere respond to erratic space weather.
    (c) List some of the adverse effects that can occur as a result of a geomagnetic storm.
16. (a) Describe the production of the aurora borealis and aurora australis.
    (b) Describe how the auroras change during a geomagnetic storm.

14.1 RADIATION PENETRATION

Aim

To distinguish between the different types of radiation according to their different penetration powers

Apparatus

radiation detector such as a Geiger–Muller tube and counter
alpha source, such as americium-241
beta source, such as strontium-90
gamma source, such as cobalt-60
sheet of paper and a sheet of lead foil

Theory

In the very early twentieth century, physicists such as Marie and Pierre Curie and Ernest Rutherford managed to identify three different types of nuclear radiation. They named these alpha (α), beta (β) and gamma (γ) rays. One of the first methods used to distinguish between them was to measure their different penetration powers — alphas were found to be poorer penetrators than betas, while gammas were much better than both.

Method

Warning

In this experiment you will be required to use radioactive isotopes. Do not handle the isotopes — you must use gloves and an appropriate handling tool. Minimise exposure to the isotopes, returning them to their box as soon as possible. Do not allow them to become wet. Do not allow any contact, directly or indirectly, with your mouth.

1. Plug the Geiger–Muller tube into the counter and connect an appropriate power supply. The counter should begin to register a slow background radiation. Let this continue for several minutes to gauge the magnitude of this background.
2. Begin with the alpha source and place it immediately in front of the tube, so that the counter indicates a high level of radiation. Now move the source away from the tube until the counter drops back to the background level. This will happen quite suddenly. Measure the distance between the source and the front of the tube, and record this information in your results table.
3. Once again, place the source immediately in front of the tube to register a high count rate. Insert a sheet of paper between the tube and the source. Repeat this with a sheet of lead foil. Record your observations in a table.
4. Repeat this whole procedure with the beta source and the gamma source, recording your results. When testing gamma's penetration of the lead foil you may wish to fold the foil to several thicknesses to better gauge its penetrating ability.

Results

Your results table should resemble that shown below.

TYPE OF RADIATION	DISTANCE PENETRATED IN AIR (cm)	PENETRATION OF PAPER	PENETRATION OF LEAD FOIL	PENETRATING POWER (HIGH/ MEDIUM/LOW)
Alpha				
Beta				
Gamma				

Questions

1. Do your results agree with the expected pattern of relative penetrating power?
2. Write down the physical nature of each of the types of radiation.
3. Can you relate particle size to penetration power? Explain.
4. What other properties of nuclear radiation can be used to distinguish them? Describe how you could do this.

GLOSSARY

A

acceleration: the time rate of change of velocity 178

accretion: the process of the growth of a body by gathering or aggregating more matter 267

active wire: one of the wires that brings electricity to a household from a power station. This wire has a voltage that varies from +340 V to −340 V with respect to the neutral wire. 130

air resistance: a force applied to an object by the air through which it is moving 203

alpha: α, helium nuclei; one of the three types of radiation 288

ammeter: an instrument used to measure the electric current in an electric circuit. An ammeter is connected into a circuit in series. 109

ampere: (A), the SI unit of electric current. An ampere is equivalent to a coulomb second^{-1}. 105

amplitude: the maximum size of the particle displacement from the undisturbed state. In relation to sound waves, the higher the amplitude, the louder the sound. 18

analogue: signals that send information over a continuous range of values 62

animal electricity: a term used by Galvani for a form of electricity that he believed was generated by animal tissues 78

attenuation: the fall off in energy that occurs as a wave passes through a medium 37

B

bandwidth: describes a series of adjacent frequencies forming a band within the spectrum 39

beats: refers to the change in volume of a sound that occurs when two sounds of slightly different frequencies occur together 23

beta: β, electrons; one of the three types of radiation 288

big bang: a term that refers to the rapid expansion of the very early universe 264

black body: a body that absorbs all radiation falling upon it. When it becomes hotter than its surroundings, it begins to radiate electromagnetic energy of its own. 276

black-body radiation: the electromagnetic radiation emitted by a black body 276

brightness: (of a star) the intensity of light as seen some distance away from it. It is the energy received per square metre per second. 274

C

centripetal acceleration: the acceleration of an object travelling in a circular path with constant speed. It is directed towards the centre of the circle. 215

centripetal force: the net force on an object travelling in a circular path at constant speed. It is directed towards the centre of the circle. 216

charge carrier: a charged particle that is free to move through a material 94

chemical electricity: the electricity produced by a chemical reaction 79

component: a part. Any vector can be resolved into a number of components. When all of the components are added together, the result is the original vector. 204

compression: a zone where the particles of the medium are pushed closer together. It is a zone of higher pressure. 16

conductor: a material that contains charge carriers 94

converging: coming together at a point (the focus) 48

cosmic background radiation: a pattern of background radiation from space that represents the afterglow of the heat of the early universe. Today it is in the form of microwaves. 262

coulomb: (C), the SI unit of electric charge 93

critical angle: the angle where total internal reflection prevents the ray from escaping from a higher optical-density medium to a lower optical-density medium 60

cycle: a complete variation in the voltage 131

D

deficiency of electrons: exists when a body has fewer electrons than protons 93

digital: communication systems that are based on signals that have two values, on or off (1s and 0s) 62

direction of a magnetic field: the direction of the force on a very small magnetic north pole placed in the field 155

displacement: a measure of the change in position of an object. It is a vector quantity. 173

diverging: lens, or mirror, from which rays appear to come from a focal point behind the lens or mirror and are spread out by the lens or mirror 49

E

earthed: when a body is connected to the Earth by a conducting path 94

echo: a repeated sound created by the reflection of sound waves from a surface 19

electric charge: a property of electrons and protons by which they exert electric forces on one another 92

electric current: the rate at which charge flows under the influence of an electric field 104

electric field: a field of force with a field strength equal to the force per unit charge at that point 98

electric field strength: $\boldsymbol{E}$, given by the formula $\boldsymbol{E} = \dfrac{\mathbf{F}}{q}$. The direction of the electric field strength is the direction of the force that acts on a positive charge placed in the field. 98

electric generator: a device in which electrical energy is produced by rotating a coil in a magnetic field 82

electric potential energy: the potential energy of an electric charge in an electric field 102

electric shock: a violent disruption of the nervous and muscular systems caused by the passage of an electric current through the body 131

electromagnetic spectrum: the full range of wavelengths of all electromagnetic waves 29

electromagnetic wave: a wave that propagates as a perpendicular electric and magnetic field. Electromagnetic waves do not require a medium for propagation. 7

electron drift: the slow movement of electrons through a conductor in the opposite direction to the electric field. This movement is superimposed on the much faster, random motion of the electrons. 106

electrostatic charge: a charge due to an excess or deficiency of electrons 93

energy: the capacity to do work. It is a scalar quantity. 226

excess of electrons: exists when a body has more electrons than protons 93

F

fibrillation: a condition in which the heart stops beating regularly and oscillates rapidly 132

focus: the point where all rays from a converging lens or mirror are concentrated. It is also the point where the rays appear to originate after passing through a diverging lens or being reflected by a diverging mirror. 48

force: a push or a pull. Force is a vector quantity. 197

free electrons: electrons in a metal that are detached from their atoms and are free to move through the metal. A metal conducts an electric current by the movement of the free electrons. 105

fuel-burning power station: a power station in which fossil fuels are burnt to provide energy 82

G

gamma: γ, high-frequency electromagnetic radiation; one of the three types of radiation 288

geocentric model: a model of the universe that has the Earth placed at its centre 253

geomagnetic storm: a period during which the Earth's magnetic field experiences unusual distortions and fluctuations in strength 299

gravitational field strength: (g), the force of gravity on a unit of mass 197

H

heliocentric model: a model of the universe that has the Sun placed at its centre 253

Hertzsprung–Russell diagram: a graph of a star's luminosity (as the vertical axis) plotted against its temperature or colour 277

Hubble's constant: a measure of the rate of expansion of the universe. It is related to the age of the universe. 262

hydro-electric power station: a power station in which water that has gained kinetic energy by flowing downhill is used to provide energy 83

I

impulse: the product of the force and the time interval over which it acts. Impulse is a vector quantity with SI units of N s. 231

induced charge: a charge produced in a body when another charged body is near it 95

induction: the production of induced charges 95

inertia: the tendency of an object to resist a change in its motion 202

instantaneous speed: the speed at a particular instant of time 177

instantaneous velocity: the velocity at a particular instant of time 177

insulated: when a body is not earthed 94

insulator: a material that does not contain charge carriers 94

interference: occurs when waves meet and interact as they pass through each other, reinforcing or cancelling at different points 20

K

kilowatt-hour: (kW-h), a unit of energy equal to the amount of energy used by a 1 kW device in 1 hour 152

kinetic energy: the energy associated with the movement of an object 227

L

lines of electric field: the lines drawn on a diagram to represent the direction and magnitude of an electric field 100

luminosity: of a star; the total energy radiated by a star per second 274

M

magnetic field: a force field surrounding a magnetic pole that exerts forces on other magnetic poles placed in the field 155

magnetosphere: the region surrounding a planet that contains its distorted magnetic field 298

main sequence: a diagonal band from the upper left corner of an H–R diagram to the lower right corner of the diagram. Most stars are in this group. 278

mechanical wave: a wave that requires the movement of particles to propagate forward 7

modulation: the process of changing the amplitude or frequency of a wave to add a signal 39

momentum: the product of the mass of an object and its velocity. It is a vector quantity. 231

N

negative charge: the type of charge on an electron 92

negatively charged: a body that has an excess of electrons 93

net force: the vector sum of the forces acting on an object 199

neutral: a body that has equal numbers of protons and electrons 93

neutral wire: one of the wires that delivers electricity to a household from a power station. This wire is maintained at earth potential. 130

newton coulomb^{-1}: ($N\ C^{-1}$), is the unit of electric field strength 99

normal: the line that is perpendicular to the reflecting surface at the point where the ray hits it 47

normal reaction: a force that acts perpendicular to a surface as a result of an object applying a force to the surface 199

nuclear fusion: the joining of nuclei to form larger ones 290

nuclear power station: a power station in which nuclear reactions provide energy 83

O

ohm: (Ω), the unit of resistance 110

oscillation: a vibration about a fixed or equilibrium point 7

overloaded circuit: one that carries a current higher than the maximum safe value for which the circuit was designed 133

P

peak voltage: the maximum voltage between the active wire and the neutral wire 131

photosphere: the visible surface of the Sun 291

pitch: directly related to the frequency of a sound. The higher the frequency of the sound, the more vibrations per second and the higher the pitch. A low-frequency sound is a low-pitched sound. 18

plasma: an extremely hot and highly ionised state of matter 265

positive charge: the type of charge on a proton 92

positively charged: a body that has a deficiency of electrons 93

potential difference: the change in potential energy per unit charge moving between two points 103

potential difference across a power supply: the number of joules of electric potential energy given to each coulomb of charge that passes through the power supply 108

potential difference across a resistor: the number of joules of electric potential energy dissipated for each coulomb of charge that passes through the resistor 107

potential energy: the energy stored in an object 230

power: the rate at which energy is transformed from one form into another 147

power supply: a source of electric potential energy 107

propagate: to transmit through space or a medium 4

R

rarefaction: a zone where the particles of the medium are spread further apart. It is a zone of lower pressure. 16

red giants: the group of stars in the upper right corner of the H–R diagram. These are cool, giant stars. 278

resistance: of a resistor, is the potential difference across the resistor divided by the current passing through the resistor 110

resistor: a conductor in which the electric potential energy of a current is converted into heat energy 106

right-hand grip rule: a rule for finding the direction of the magnetic field surrounding an electric current 157

S

scalar: a quantity that specifies size (magnitude) but not direction 173

short circuit: where an active wire comes in contact with the neutral wire or is earthed 133

sine wave: the curve that results when a plot is made of $y = \sin x$ 3

solar cycle: an 11-year cyclical pattern of increasing and decreasing frequency of sunspots, flares, prominences and coronal mass ejections 295

solar wind: an outflow of low-density plasma from the corona of the Sun 292

speed: a measure of the time rate at which an object moves over a distance. It is a scalar quantity. 174

sunspot: a dark spot on the surface of the Sun, representing a region of intense magnetic activity and lower temperature 294

sunspot cycle: a cyclical pattern of increasing and decreasing numbers of sunspots 295

superposition: the adding of two or more waves 21

T

timbre: the quality of a sound that depends on the way in which a number of different pure sounds have combined 23

U

uniform electric field: an electric field with the same magnitude and direction at all points 101

V

Van Allen belt: either of two doughnut-shaped zones of radiation (ions) that wrap around the Earth 299

vector: a quantity that specifies size (magnitude) and direction 173

velocity: a measure of the time rate of displacement, or the time rate of change in position. It is a vector quantity. 174

volt: (V), the SI unit of potential difference 103

voltage: another name for potential difference 103

voltaic pile: an assembly of large numbers of alternate zinc and brass discs separated by cardboard discs soaked in salt solution 79

voltmeter: an instrument used to measure the potential difference across a component in an electric circuit. A voltmeter is connected into a circuit in parallel. 109

W

watt: (W), the SI unit of power 147

weight: the force applied to an object due to gravitational attraction 197

white dwarfs: the group of stars close to the lower left corner of the H–R diagram. These are exceptionally small, hot stars. 278

work: done when an object moves in the direction of a force applied to it. The amount of work done is the product of the magnitude of the force and the displacement of the object in the direction of the force. Work is a scalar quantity. 226

APPENDIX 1: Formulae and data sheet

Formulae

The world communicates

$v = f\lambda$

$I \propto \frac{1}{d^2}$

$\frac{v_1}{v_2} = \frac{\sin i}{\sin r}$

Electrical energy in the home

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

$R = \frac{V}{I}$

$P = VI$

Energy $= VIt$

Moving about

$v_{av} = \frac{\Delta r}{\Delta t}$

$a_{av} = \frac{\Delta v}{\Delta t}$ therefore $a_{av} = \frac{v - u}{t}$

$\Sigma F = ma$

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

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

$W = Fs$

$p = mv$

Impulse $= Ft$

The cosmic engine

brightness $= \frac{\text{luminosity}}{4\pi r^2}$

$\lambda_{\max} T = W$

$v = H_0 D$

Data sheet: numerical values of constants

Charge on the electron, q_e	-1.602×10^{-19} C
Mass of electron, m_e	9.109×10^{-31} kg
Mass of neutron, m_n	1.675×10^{-27} kg
Mass of proton, m_p	1.673×10^{-27} kg
Speed of sound in air	340 m s^{-1}
Earth's gravitational acceleration, g	9.8 m s^{-2}
Speed of light, c	3.00×10^{8} m s^{-1}
Magnetic force constant $\left(k \equiv \frac{\mu_0}{2\pi}\right)$	2.0×10^{-7} N A^{-2}
Universal gravitational constant, G	6.67×10^{-11} N m^2 kg^{-2}
Mass of Earth	6.0×10^{24} kg
Planck's constant, h	6.626×10^{-34} J s
Rydberg's constant, R_H (hydrogen)	1.097×10^{7} m^{-1}
Atomic mass unit, u	1.661×10^{-27} kg
	$931.5 \frac{\text{MeV}}{c^2}$
1 eV	1.602×10^{-19} J
Density of water, ρ	1.00×10^{3} kg m^{-3}
Specific heat capacity of water	4.18×10^{3} J kg^{-1} K^{-1}

APPENDIX 2: Periodic table

Key: Atomic number / Symbol / Atomic weight / Name

Period	Group 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1 H 1.008 Hydrogen																	2 He 4.003 Helium
2	3 Li 6.941 Lithium	4 Be 9.012 Beryllium											5 B 10.81 Boron	6 C 12.01 Carbon	7 N 14.01 Nitrogen	8 O 16.00 Oxygen	9 F 19.00 Fluorine	10 Ne 20.18 Neon
3	11 Na 22.99 Sodium	12 Mg 24.31 Magnesium											13 Al 26.98 Aluminium	14 Si 28.09 Silicon	15 P 30.97 Phosphorus	16 S 32.07 Sulfur	17 Cl 35.45 Chlorine	18 Ar 39.95 Argon
4	19 K 39.10 Potassium	20 Ca 40.08 Calcium	21 Sc 44.96 Scandium	22 Ti 47.87 Titanium	23 V 50.94 Vanadium	24 Cr 52.00 Chromium	25 Mn 54.94 Manganese	26 Fe 55.85 Iron	27 Co 58.93 Cobalt	28 Ni 58.69 Nickel	29 Cu 63.55 Copper	30 Zn 65.39 Zinc	31 Ga 69.72 Gallium	32 Ge 72.61 Germanium	33 As 74.92 Arsenic	34 Se 78.96 Selenium	35 Br 79.90 Bromine	36 Kr 83.80 Krypton
5	37 Rb 85.47 Rubidium	38 Sr 87.62 Strontium	39 Y 88.91 Yttrium	40 Zr 91.22 Zirconium	41 Nb 92.91 Niobium	42 Mo 95.94 Molybdenum	43 Tc [98] Technetium	44 Ru 101.1 Ruthenium	45 Rh 102.9 Rhodium	46 Pd 106.4 Palladium	47 Ag 107.9 Silver	48 Cd 112.4 Cadmium	49 In 114.8 Indium	50 Sn 118.7 Tin	51 Sb 121.8 Antimony	52 Te 127.6 Tellurium	53 I 126.9 Iodine	54 Xe 131.3 Xenon
6	55 Cs 132.9 Caesium	56 Ba 137.3 Barium	57–71 * Lanthanides	72 Hf 178.5 Hafnium	73 Ta 180.9 Tantalum	74 W 183.8 Tungsten	75 Re 186.2 Rhenium	76 Os 190.2 Osmium	77 Ir 192.2 Iridium	78 Pt 195.1 Platinum	79 Au 197.0 Gold	80 Hg 200.6 Mercury	81 Tl 204.4 Thallium	82 Pb 207.2 Lead	83 Bi 209.0 Bismuth	84 Po [209] Polonium	85 At [210] Astatine	86 Rn [222] Radon
7	87 Fr [223] Francium	88 Ra [226] Radium	89–103 ** Actinides	104 Rf [261] Rutherfordium	105 Db [262] Dubnium	106 Sg [266] Seaborgium	107 Bh [264] Bohrium	108 Hs [265] Hassium	109 Mt [268] Meitnerium	110 Ds [271] Darmstadtium	111 Rg [280] Roentgenium	112 Uub [272] Ununbium	113	114 Uuq [285] Ununquadium	115	116 Uuh [289] Ununhexium	117	118 Uuo [294] Ununoctium

*Lanthanide series

57 La 138.9 Lanthanum	58 Ce 140.1 Cerium	59 Pr 140.9 Praseodymium	60 Nd 144.2 Neodymium	61 Pm [145] Promethium	62 Sm 150.4 Samarium	63 Eu 152.0 Europium	64 Gd 157.3 Gadolinium	65 Tb 158.9 Terbium	66 Dy 162.5 Dysprosium	67 Ho 164.9 Holmium	68 Er 167.3 Erbium	69 Tm 168.9 Thulium	70 Yb 173.0 Ytterbium	71 Lu 175.0 Lutetium

**Actinide series

89 Ac [227] Actinium	90 Th 232.0 Thorium	91 Pa 231.0 Protactinium	92 U 238.0 Uranium	93 Np [237] Neptunium	94 Pu [244] Plutonium	95 Am [243] Americium	96 Cm [247] Curium	97 Bk [247] Berkelium	98 Cf [251] Californium	99 Es [252] Einsteinium	100 Fm [257] Fermium	101 Md [258] Mendelevium	102 No [259] Nobelium	103 Lr [262] Lawrencium

- For elements with no stable nuclides, the mass of the longest living isotope is given in square brackets.
- The atomic weights of Np and Tc are given for the isotopes ^{237}Np and ^{99}Tc.

APPENDIX 3: Key words for examination questions

HSC syllabus documents and examination questions use the following key words that state what students are expected to be able to do.

Account	Account for: state reasons for, report on. Give an account of: narrate a series of events or transactions
Analyse	Identify components and the relationship between them; draw out and relate implications
Apply	Use, utilise, employ in a particular situation
Appreciate	Make a judgement about the value of
Assess	Make a judgement of value, quality, outcomes, results or size
Calculate	Ascertain/determine from given facts, figures or information
Clarify	Make clear or plain
Classify	Arrange or include in classes/categories
Compare	Show how things are similar or different
Construct	Make; build; put together items or arguments
Contrast	Show how things are different or opposite
Critically (analyse/evaluate)	Add a degree or level of accuracy, depth, knowledge and understanding, logic, questioning, reflection and quality to (analysis/evaluation)
Deduce	Draw conclusions
Define	State meaning and identify essential qualities
Demonstrate	Show by example
Describe	Provide characteristics and features
Discuss	Identify issues and provide points for and/or against
Distinguish	Recognise or note/indicate as being distinct or different from; to note differences between
Evaluate	Make a judgement based on criteria; determine the value of
Examine	Inquire into
Explain	Relate cause and effect; make the relationships between things evident; provide why and/or how
Extract	Choose relevant and/or appropriate details
Extrapolate	Infer from what is known
Identify	Recognise and name
Interpret	Draw meaning from
Investigate	Plan, inquire into and draw conclusions about
Justify	Support an argument or conclusion
Outline	Sketch in general terms; indicate the main features of
Predict	Suggest what may happen based on available information
Propose	Put forward (for example, a point of view, idea, argument, suggestion) for consideration or action
Recall	Present remembered ideas, facts or experiences
Recommend	Provide reasons in favour
Recount	Retell a series of events
Summarise	Express, concisely, the relevant details
Synthesise	Put together various elements to make a whole

ANSWERS TO NUMERICAL QUESTIONS

CHAPTER 1

7. $f = 7.317 \times 10^{14}$ Hz
8. 0.8 Hz
11. $\lambda = 1.29$ m

CHAPTER 2

1. $\lambda = 0.644\ m$
8. 384 vibration cycles per second (or 384 Hz)

CHAPTER 4

2. 40°
14. $v_1 = 2.041 \times 10^8$ m s^{-1}
15. $r = 19.5°$
16. sin $i_c = 41.8°$

CHAPTER 6

3. 1.60×10^{-19} C
4. 1.25×10^{13}
5. (a) $+5.44 \times 10^{-14}$ C
   (b) -5.44×10^{-14} C
7. 2.17×10^2 N C^{-1} east
8. 3.00×10^{-1} N C^{-1} north
9. (a) 6.88×10^{-17} N west
   (b) 6.88×10^{-17} N east
   (c) 9.89×10^{-2} N east
   (d) 2.79×10^{-1} N west
13. (a) 2.00×10^2 V
    (b) 5.20×10^{-1} J
16. (a) 6.00 C
    (b) 3.75×10^{19}
17. 1.40 A
18. (a) 8.96×10^{-1} C
    (b) 3.58×10^{-1} A
19. (a) (i) 5.00 C
        (ii) 3.00×10^2 C
    (b) 3.13×10^{19}, 1.88×10^{21}
22. 2.00×10^1 V
23. 3.84×10^2 J
25. 5.00 J
27. 8.00 Ω
28. 6.40 V
29. (a) 5.00 Ω
    (b) 72.0 V
    (c) 0.500 A
    (d) 0.167 Ω
    (e) 3.00 A
    (f) 10.0 V
30. (a) 24.0 J
    (b) 2.00 V
    (c) 7.00 C
    (d) 4.00 V
    (e) 0.125 C
    (f) 32.0 J
31. b, a, d, c

CHAPTER 7

2. (a) 5.00 A
   (b) 5.00 A
3. (a) 2.00 A
   (b) R_2
4. 8.00 A
5. (a) 3.00 V
   (b) R_1
6. 4.0 V
7. (a) 20.0 V
   (b) 20.0 V
8. (a) 3.00 A
   (b) 3.00 A
   (c) 12.0 V
   (d) 48.0 V
9. (a) 20.0 V
   (b) 40.0 V
   (c) 4.00 A
   (d) 10.0 Ω
10. 4.00 A
11. (a) 24.0 V
    (b) 24.0 V
    (c) 6.00 A
    (d) 24.0 V
    (e) 8.00 A
12. (a) 3.00 A
    (b) 5.00 A
    (c) 8.00 A
13. (a) 36.0 Ω
    (b) 2.00 A
    (c) 24.0 V
14. 3.00 Ω
15. (a) 28.0 Ω
    (b) 3.00 Ω
    (c) 39.0 Ω
    (d) 4.13 Ω
    (e) 90.0 Ω
    (f) 0.400 Ω
16. (a) (i) 0.500 A, 0.500 A
        (ii) 2.50 V, 3.50 V
    (b) (i) 2.00 A, 2.00 A, 2.00 A
        (ii) 20.0 V, 40.0 V, 80.0 V
    (c) (i) 2.5 A, 2.5 A
        (ii) 5.9 V, 9.1 V
    (d) (i) 2.03 A, 2.03 A, 2.03 A
        (ii) 22.7 V, 41.4 V, 63.9 V
    (e) (i) 6.00 A, 4.00 A
        (ii) 12.0 V, 12.0 V
    (f) (i) 4.00 A, 2.00 A, 0.800 A
        (ii) 48.0 V, 48.0 V, 48.0 V
    (g) (i) 7.23 A, 4.88 A
        (ii) 125 V, 125 V
    (h) (i) 12.1 A, 4.28 A, 7.13 A
        (ii) 30.5 V, 30.5 V, 30.5 V

CHAPTER 8

1. 3.00×10^1 W
2. 3.00×10^4 J
3. 2.00×10^1 s

4. 6.0 kW-h
5. (a) 6.0×10^5 J
   (b) 1.25×10^1 W
   (c) 2.2×10^8 J (62 kW-h)
   (d) 3.3×10^{-4} s
   (e) 6.86 W
   (f) 12.5 s
   (g) 4200 J
   (h) 2.0×10^2 W
   (i) 5.40 kW-h
   (j) 2.00 kW
   (k) 3.30 h
6. (a) 2.10 V
   (b) 2.00 A
   (c) 10.0 W
   (d) 0.33 A
   (e) 0.50 V
   (f) 62.5 W
7. (a) 2.00 A
   (b) 2.50 s
   (c) 1.00 V
   (d) 215 J
   (e) 5.63 A
   (f) 159 J
   (g) 1.80×10^3 s
   (h) 34.8 V
8. (a) 4.2 A , 58 Ω
   (b) 3.58×10^2 W, 35.0 Ω
   (c) 190 W, 96 V
   (d) 130 W, 8.0 A
   (e) 670 V, 4400 Ω
   (f) 2.50 A, 80.0 Ω

CHAPTER 9

1. b, d
2. (a) 17 km
   (b) 13 km S 67° W
   (c) 24 m s^{-1}
   (d) 18 m s^{-1} S 67° W
3. (a) 14.67 m s^{-1}
   (b) 2 h 56 min 6 s
   (c) 1.9 h
   (d) (i) 78 km h^{-1} (ii) 0 km h^{-1}
5 (b) 89 km h^{-1}
7. 50 s
8. (a) 1.8 s
   (b) 5.0 s
9. (a) (i) −40 km h^{-1}
   (ii) 40 km h^{-1} south or −40 km h^{-1} north
   (b) (i) −20 m s^{-1}
   (ii) −20 m s^{-1} in original direction
   (c) (i) +5 m s^{-1}
   (ii) −55 m s^{-1} in original direction
11. (a) 4.0 m s^{-2} north
   (b) 10 m s^{-2} N 37° E
   (c) 2.8 m s^{-2} south-east
12. (a) 2.7 s
   (b) 27 m s^{-1}
13. (a) 12 m s^{-1}
   (b) −6.0 m s^{-2}
14. (a) 3.45 s
   (b) −8.05 m s^{-2}
15. (a) 3.2 s
   (b) 3.8 m s^{-2} down the slope
   (c) 1.6 s
16. (a) B, C
   (b) B, D
   (c) A, E
   (d) A, E
   (e) D
17. (a) 10 m south
   (b) 5.0 m s^{-2} north
   (c) 3.0 m s^{-2} south
   (d) 6.1 m s^{-1} south
18. (a) B
   (b) A, D, E
   (c) 40 s
   (d) 20 m north
   (e) 260 m
   (f) D
   (g) E
   (h) 3.3 m s^{-1}
   (i) 6.0 m s^{-1} south
   (j) approx. 3 m s^{-1} north
20. (a) B, D, F
   (b) +20 m
   (c) 0.25 m s^{-1}
   (d) 30 s
   (e) It didn't
   (f) C, G
   (g) 20–30 s, 50–55 s, 70–80 s
   (i) 0.20 m s^{-2}
   (j) 0.050 m s^{-2}
22. (a) The jet skier after 8.0 s
   (b) (i) 21 m s^{-1} (ii) 33 m s^{-1}
23. 3.7 h or 3 h 41 min
24. (a) 3.0 s
   (b) 2.5 m s^{-2}
   (c) 10 s
   (d) 80 m
25. 995 m

CHAPTER 10

2. b, c, e
4. (a) 1.4×10^4 N
   (b) 5.0×10^3 N
   (c) 1400 kg
5. If your mass is *m* kg;
   (a) 9.8*m* N
   (b) 3.6*m* N
   (c) *m* kg
6. The beam balance
7. (a) 3 N east
   (b) 1.4×10^2 N east
8. (a) 346 N east
   (b) 53.6 N east
9. (a) (i) equal to

(ii) equal to
(iii) greater than
(iv) less than

10. (b) zero
16. (a) 1.8×10^2 N
(b) 100 N
(c) zero
17. (b) down the hill
(c) weight
18. (a) zero
(b) 3.9×10^2 N (392 N)
(c) 4.0×10^2 N (402 N)
(d) 6.8×10^2 N
19. (a) down the slope
(c) 2.9×10^2 N (294 N)
(d) 2.9×10^2 N (286 N)
20. (a) zero
(b) 9.8 m s^{-2}
(c) 4.9 N down
21. (a) 6.4×10^3 N
(b) zero
(c) 6.6×10^6 N
22. (a) 7500 N
(b) 6.3 m s^{-2}
(c) 31 m s^{-1}
(d) 78 m
23. (a) 9.6×10^6 N
(b) 3.8×10^2 m (375 m)
24. (a) 823 N
(b) 6.6×10^2 N (655 N)
(c) 9.9×10^2 N
25. 5.8 m s^{-2}
26. (a) 42 N
(b) 19°
31. (a) 2.0 m s^{-2} to the right
(b) 6.0 N
(c) 8.0 N to the right
(d) 3.5 m s^{-2} to the right
32. (a) 4.0 m s^{-2} to the right
(b) 160 N to the right
(c) 240 N to the left
(d) 240 N to the right
33. 0.86 m s^{-2} to the right
34. (a) 12 m s^{-2}
(b) 1.8×10^4 N
(c) 7.2×10^2 N
35. (a) 3.2×10^3 N
(b) 3.2×10^3 N towards the centre of the racetrack
(c) 6.3×10^2 m (6.25×10^2 m)
37. (a) 7840 N towards the centre of the track
(b) 56 m s^{-1}

CHAPTER 11

3. 59 J
4. zero
5. (a) approx. 1×10^5 J
(b) approx. 4×10^1 J
(c) approx. 3×10^3 J
(d) approx. 5×10^{-9} J
6. (a) 1.8×10^5 J
(b) 1.8×10^5 J
(c) 17 m s^{-1}
7 4.6×10^3 N
9. (a) 8×10^2 kg m s^{-1}
(b) 1×10^4 kg m s^{-1}
(c) 6×10^2 N s
(d) 4 N s
(e) 10 kg m s^{-1}
10. (a) 1500 N s due east
(b) 1500 kg m s^{-1} due east
(c) 750 N
11. (a) 60 N s
(b) 90 kg m s^{-1}
12. (a) 2.3×10^4 N s opposite to the initial direction of the car
(b) 2.9×10^5 N opposite to the initial direction of the car
(c) 2.1×10^2 m s^{-2}
13. (a) approx. 160 N s
(b) 2.7 m s^{-1}
15. (a) 240 N s upwards
(b) 3.1×10^3 N upwards
19. (a) 3.0 kg m s^{-1} opposite to the initial direction of the toy car
(b) 3.0 N s in the initial direction of the car
20. (a) 0.30 m s^{-1}
(b) 0.60 m s^{-1}
21. (a) 40 kg
(b) 60 N s
(c) 60 N s
(d) zero
22. (a) 1.7 m s^{-1}
(b) 120 N s
(c) 120 kg m s^{-1}
(d) 120 N s
(f) 0.92 m s^{-1}
23. (a) 15 m s^{-1}
(b) 1.1×10^4 N s in the initial direction of the car
(c) 420 N s opposite to the initial direction of the car
(d) 1.1×10^5 N
24. (a) 2.9 m s^{-1} east
(b) The car
(c) They experience the same
(d) They experience the same
25. 2.9 m s^{-1}

CHAPTER 13

2. (a) Four times brighter
(b) Sixteen times brighter
4. Similar brightness; the more distant star 16 times more luminous than the other.
5. Rigel (1800 times more luminous but approximately 150 times further away) is one-sixth the brightness of Sirius.
6. Fomalhaut: 0.002 L_0 watts per square light year
Bellatrix: 0.001 L_0 watts per square light year.
Hence, Fomalhaut is the brighter of the two stars.

INDEX