PEARSON

PHYSICS

WESTERN AUSTRALIA

STUDENT BOOK

11

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

www.pearson.com.au

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

Publisher: Malcolm Parsons
Project Manager: Shelly Wang
Production Manager: Anji Bignell
Development Editor: Vicky Chadfield
Designer and cover designer: Anne Donald
Rights & Permissions Editor: Amirah Fatin
Editor: Sam Trafford
Senior Publishing Services Analyst: Rob Curulli
Proofreader: Jane Fitzpatrick
Indexer: Max McMaster
Illustrator/s: DiacriTech
Printed in ** by **

National Library of Australia Cataloguing-in-Publication entry
Pearson physics 11 : Western Australia / Greg Moran (coordinating author).

ISBN: 9781488617713 (pbk.)

Target Audience: For secondary school age.

Subjects: Physics—Study and teaching (Secondary)—Western Australia.

Physics—Textbooks.

Pearson Australia Group Pty Ltd ABN 40 004 245 943

Acknowledgements
All third party material contained in this sample is for placement only, permission will be sought and all material appropriately acknowledged in the finished product.

We would like to thank the following for permission to reproduce copyright material.

WACE Physics ATAR Course Year 11 Syllabus © School Curriculum and Standards Authority, Government of Western Australia, 2014

Every effort has been made to trace and acknowledge copyright. However, if any infringement has occurred, the publishers tender their apologies and invite the copyright holders to contact them.

Some of the images used in *Pearson Physics 11 Western Australia Student Book* might have associations with deceased Indigenous Australians. Please be aware that these images might cause sadness or distress in Aboriginal or Torres Strait Islander communities.

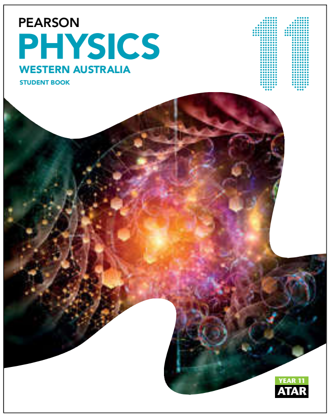
PEARSON

PHYSICS

WESTERN AUSTRALIA

11

Writing and Development Team



Greg Moran
Head of Science,
Past President STAWA
Coordinating author

Doug Bail
Education Consultant,
Past Head of Science
Author

Stephen Brown
Teacher
Reviewer

Subrat Das
Educator
Reviewer

Tracey Fisher
Teacher,
Former Lecturer, Research Scientist
Author and reviewer

Geoff Lewis
Head of Science, Past President STAWA
Author

Elke McKay
Teacher
Author and reviewer

Gregory White
Physics Educator
Author

Daniela Nardelli
Teacher and Physics Consultant
Contributing Author

Craig Tilley
Science Writer
Contributing Author

Alistair Harkness
Teacher
Contributing Author

Jack Jurica
Teacher
Contributing Author

Contents

UNIT 1 Thermal, nuclear and electrical physics

AREA OF STUDY 1

Heating Processes

CHAPTER 1 **Heating processes** 3

1.1 Heat and temperature 4

1.2 Specific heat capacity 11

1.3 Latent heat 15

1.4 Heating and cooling 21

Chapter review 25

CHAPTER 2 **Moving heat around** 27

2.1 Work and efficiency 28

2.2 Conduction 32

2.3 Convection 36

2.4 Radiation 39

Chapter review 43

AREA OF STUDY 2

Ionising Radiation and Nuclear Reactions

CHAPTER 3 **Particles in the nucleus** 45

3.1 Atoms, isotopes and radioisotopes 46

3.2 Radioactivity 55

3.3 Properties of alpha, beta and gamma radiation 62

3.4 Half-life and decay series 68

3.5 Radiation doses and effects on humans 75

Chapter review 85

CHAPTER 4 **Fission and fusion** 87

4.1 Nuclear fission and energy 88

4.2 Chain reactions and nuclear reactors 96

4.3 Nuclear fusion 104

Chapter review 111

AREA OF STUDY 3

Electrical Circuits

CHAPTER 5 **Electrical physics** 113

5.1 Behaviour of charged particles 114

5.2 Energy in electric circuits 119

5.3 Electric current and circuits 124

5.4 Resistance 132

5.5 Series and parallel circuits 142

5.6 Electrical safety 157

Chapter review 164

Unit 1 Review 167

UNIT 2 Linear motion and waves

AREA OF STUDY 4

Linear Motion and Force

CHAPTER 6 **Scalars and vectors** **173**

6.1 Scalars and vectors 174

6.2 Adding vectors in one and two dimensions 181

6.3 Subtracting vectors in one and two dimensions 187

6.4 Vector components 193

Chapter review 196

CHAPTER 7 **Linear motion** **197**

7.1 Displacement, speed and velocity 198

7.2 Acceleration 208

7.3 Graphing position, velocity and acceleration over time 213

7.4 Equations for uniform acceleration 226

7.5 Vertical motion 232

Chapter review 239

CHAPTER 8 **Momentum and forces** **243**

8.1 Momentum and conservation of momentum 244

8.2 Change in momentum and impulse 251

8.3 Newton's first law 254

8.4 Newton's second law 263

8.5 Newton's third law 271

8.6 Impulse and force 277

8.7 Mass and weight 287

Chapter review 291

AREA OF STUDY 5

Waves

CHAPTER 9 **Work, energy and power** **293**

9.1 Energy and work 294

9.2 Kinetic energy 304

9.3 Elastic and inelastic collisions 309

9.4 Gravitational potential energy 313

9.5 Law of conservation of energy 318

9.6 Power 327

Chapter review 330

CHAPTER 10 **The nature of waves** **332**

10.1 Longitudinal and transverse waves 334

10.2 Representing waves 339

10.3 Wave behaviours—reflection, refraction and diffraction 348

10.4 Wave interactions—superposition, interference and resonance 361

10.5 Standing waves and harmonics 367

10.6 Wave intensity and applications of wave properties 380

Chapter review 388

AREA OF STUDY 6

Practical Investigation

CHAPTER 11 **Practical investigation** **390**

11.1 Designing and planning the investigation 392

11.2 Conducting investigations and recording and presenting data 400

11.3 Discussing investigations and drawing evidence-based conclusions 407

Chapter review 413

Unit 2 Review **414**

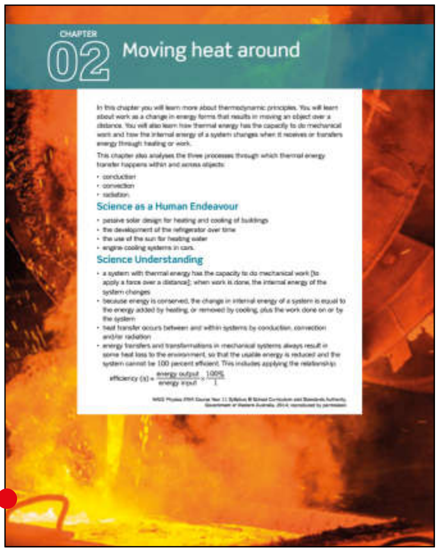
GLOSSARY **451**

INDEX **XXX**

How to use this book

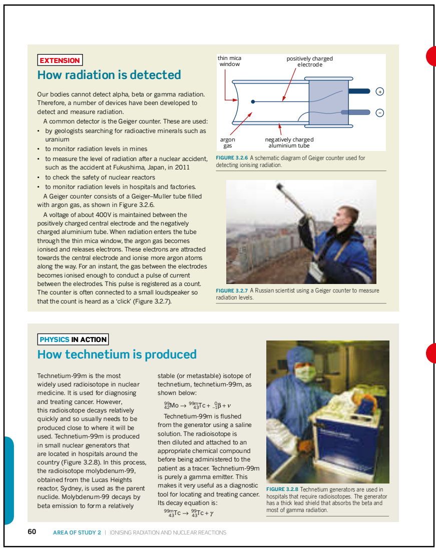
Pearson Physics 11 Western Australia

Pearson Physics 11 Western Australia has been written to the *WACE Physics ATAR Course, Year 12 Syllabus 2016*. Each chapter is clearly divided into manageable sections of work. Best practice literacy and instructional design are combined with high quality, relevant photos and illustrations. Explore how to use this book below.



Chapter opening page

The chapter opening page links the syllabus to the chapter content. Science Understanding and Science as a Human Endeavour addressed in the chapter is clearly listed.

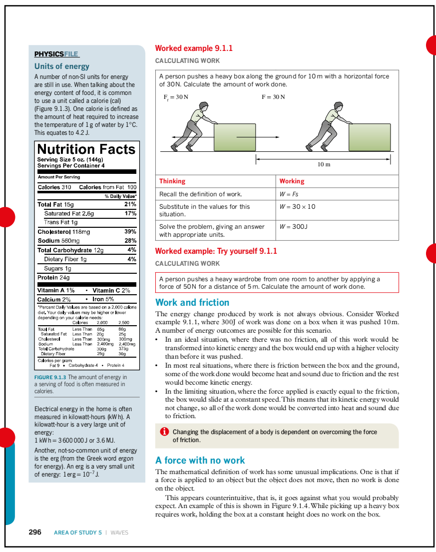


EXTENSION

Extension boxes include material that goes beyond the core content of the syllabus. They are intended for students who wish to expand their depth of understanding in a particular area.

PHYSICS IN ACTION

Physics in Action boxes place physics in an applied situation or relevant context and encourages students to think about the development of physics and its use and influence of physics in society.



PHYSICSFILE

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

Highlight box

Focuses students' attention on important information such as key definitions, formulae and summary points.

Worked examples

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

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

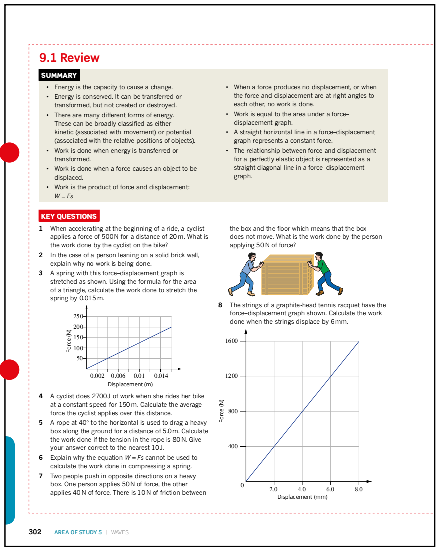
Fully worked solutions to all Try Yourself: Worked examples are available on *Pearson Physics 12 Western Australia Teacher Reader+*

Section summary

Each section includes a summary to assist students consolidate key points and concepts.

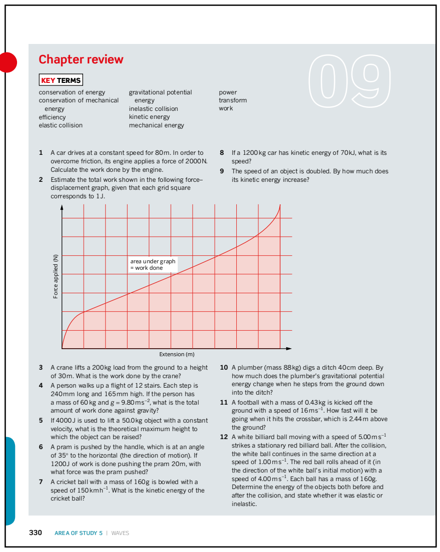
Section review questions

Each section finishes with questions to test students' understanding and ability to recall the key concepts of the section



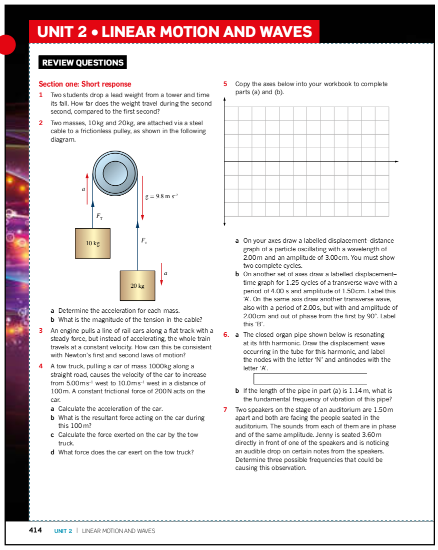
Chapter review

Each chapter finishes with a set of higher order questions to test students' ability to apply the knowledge gained from the chapter.



Unit Review

Each unit finishes with a comprehensive set of exam-style questions, that assist students draw together their knowledge and understanding and apply it to this style of questions.



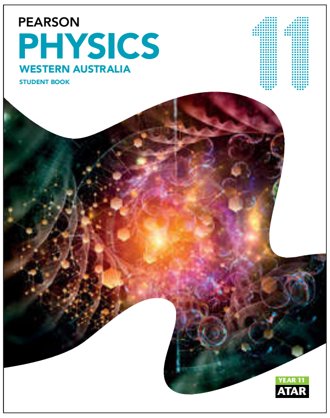
Glossary

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

Answers

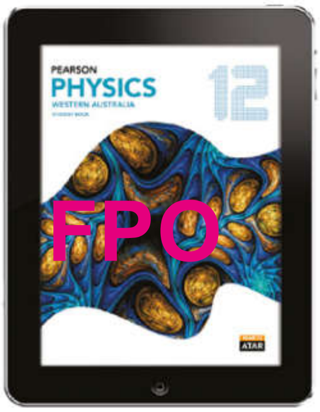
Numerical answers and key short response answers are included at the back of the book. Comprehensive answers and fully worked solutions for all section review questions, Try yourself: Worked examples, chapter review questions and Unit review questions are provided via *Pearson Physics 12 Western Australia Teacher Reader+*

Pearson Physics 11 Western Australia



Student Book

Pearson Physics 11 Western Australia has been written to fully align with the *WACE Physics ATAR Course, Year 12 Syllabus 2016.* The series includes the very latest developments and applications of Physics and incorporates best practice literacy and instructional design to ensure the content and concepts are fully accessible to all students.



Student Reader+.

Pearson Student Reader+ lets you use your Student Book online or offline on any device. Pearson Reader+ includes interactive activities to enhance learning and test understanding.

Teacher Reader+.

Pearson Teacher Reader+ provides comprehensive teacher support including teaching programs, fully worked solutions to all questions and chapter tests and practice exams. Pearson Teacher Reader+ also includes all Pearson Student Reader+ assets. A hardcopy option of selected resources is also available.

PearsonDigital

Browse and buy at **pearson.com.au**
Access the ProductLink at **pearsonplaces.com.au**

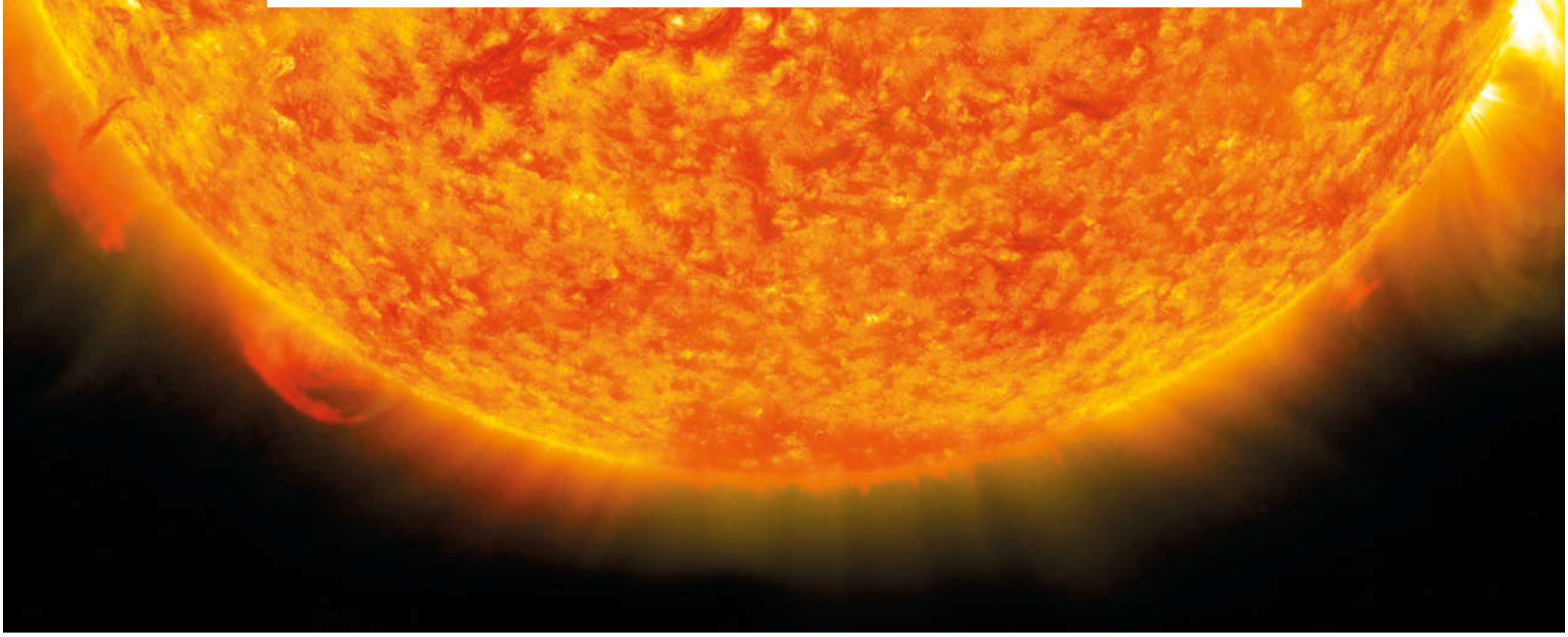
UNIT 1 Thermal, nuclear and electrical physics

An understanding of heating processes, nuclear reactions and electricity is essential to appreciate how global energy needs are met. In this unit, students explore the ways physics is used to describe, explain and predict the energy transfers and transformations that are pivotal to modern industrial societies. Students investigate heating processes, apply the nuclear model of the atom to investigate radioactivity, and learn how nuclear reactions convert mass into energy. They examine the movement of electrical charge in circuits and use this to analyse, explain and predict electrical phenomena.

Learning outcomes

By the end of this unit, students:

- understand how the kinetic particle model and thermodynamics concepts describe and explain heating processes
- understand how the nuclear model of the atom explains radioactivity, fission, fusion and the properties of radioactive nuclides
- understand how charge is involved in the transfer and transformation of energy in electrical circuits
- understand how scientific models and theories have developed and are applied to improve existing, and develop new, technologies
- use science inquiry skills to design, conduct and analyse safe and effective investigations into heating processes, nuclear physics and electrical circuits, and to communicate methods and findings
- use algebraic and graphical representations to calculate, analyse and predict measurable quantities associated with heating processes, nuclear reactions and electrical circuits
- evaluate, with reference to empirical evidence, claims about heating processes, nuclear reactions and electrical technologies
- communicate physics understanding using qualitative and quantitative representations in appropriate modes and genres



CHAPTER 01

Heating processes

Thermal energy is part of everyday experience. It's the thermal energy from the Sun that makes your world habitable. Humans can thrive in the climatic extremes of the Earth, from the outback deserts to ski slopes in winter. In this chapter, the nature of thermal energy is explored. Specifically, by the end of this section, you will have covered material relating to heating processes in the following areas:

- kinetic particle theory
- temperature
- internal energy
- specific heat capacity
- latent heat
- thermal equilibrium.

Science as a Human Endeavour

- passive solar design for heating and cooling of buildings
- the development of the refrigerator over time
- the use of the sun for heating water
- engine cooling systems in cars

Science Understanding

- the kinetic particle model describes matter as consisting of particles in constant motion, except at absolute zero
- all substances have internal energy due to the motion and separation of their particles
- temperature is a measure of the average kinetic energy of particles within a system
- provided a substance does not change state, its temperature change is proportional to the amount of energy added to or removed from the substance; the constant of proportionality describes the heat capacity of the substance.
  This includes applying the relationship:

  $$Q = mc\Delta T$$

- change of state involves separating particles which exert attractive forces on each other; latent heat is the energy required to be added to or removed from a system to change the state of the system.
  This includes applying the relationship:

  $$Q = mL$$

- two systems in contact transfer energy between particles so that eventually the systems reach the same temperature; that is, they are in thermal equilibrium. This may involve changes of state as well as changes in temperature.

1.1 Heat and temperature

In the sixteenth century, Sir Francis Bacon, an English essayist and philosopher, proposed a radical idea: that heat is motion. He went on to write that heat is the rapid vibration of tiny particles within every substance. At the time, his ideas were dismissed because the nature of particles wasn't fully understood. An opposing theory at the time was that heat was related to the movement of a fluid called 'caloric' that filled the spaces within a substance.

Today, it is understood that all matter is made up of small particles (atoms or molecules). Using this knowledge, it is possible to look more closely at what happens during heating processes.

This section starts by looking at the kinetic particle model, which states that the small particles (atoms or molecules) that make up all matter have kinetic energy. Therefore all particles are in constant motion, even in extremely cold solids. It was thought centuries ago that if a material was continually made cooler, there would be a point at which the particles would eventually stop moving. This coldest possible temperature is called **absolute zero** and will be discussed later in this section.

KINETIC PARTICLE MODEL

Some philosophers of the Middle Ages believed that heat was a fluid that filled the spaces between the particles of a substance and flowed from one substance to another. This is known as the 'caloric' theory. When caloric flowed from one substance into another, the first object cooled down and the second object heated up. Many attempts were made to detect caloric, but none were successful. It was assumed that caloric had no mass, odour, taste or colour. Scientists now know that caloric simply doesn't exist.

The best understanding of the behaviour of matter today depends on a model called the **kinetic particle model** (kinetic theory). A model is a representation that describes or explains the workings within an object, system or idea. This will generally include making some assumptions. The assumptions behind the kinetic particle model are:

- All matter is made up of many very small particles (atoms or molecules).
- The particles are in constant motion.
- Overall, no kinetic energy is lost or gained during elastic collisions between particles.
- There are forces of attraction and repulsion between the particles in a material.
- The distances between particles in a gas are large compared with the size of the particles.

The kinetic theory applies to all states (or phases) of matter including the three you most commonly come across in your everyday activities: solids, liquids and gases.

Solids

Within a solid, the particles must be exerting attractive forces or bonds on each other for the matter to hold together in its fixed shape. There must also be repulsive forces, without which the attractive forces would cause the solid to collapse. In a solid, the attractive and repulsive forces hold these particles in fixed positions, usually in a regular arrangement or lattice (see Figure 1.1.1(a)). But the particles in a solid are not completely still; they vibrate around average positions. The forces on individual particles are sometimes predominantly attractive and sometimes repulsive, depending on their exact position relative to neighbouring particles.

Liquids

Within a liquid, there is still a balance of attractive and repulsive forces. Compared with a solid, the particles in a liquid have more freedom to move around each other and will therefore take the shape of the container (see Figure 1.1.1(b)). Generally, the liquid takes up a slightly greater volume than it would in the solid state. Particles collide but remain attracted to each other, so the liquid remains within a fixed volume but with no fixed shape.

Gases

In a gas, particles are in constant, random motion, colliding with each other and the walls of the container. The particles move rapidly in every direction, quickly filling the volume of any container, and occasionally colliding with each other (see Figure 1.1.1(c)). A gas has no fixed volume. The particle speeds are high enough that, when the particles collide, the attractive forces are not strong enough to keep the particles close together. The repulsive forces cause the particles to separate and move off in other directions.

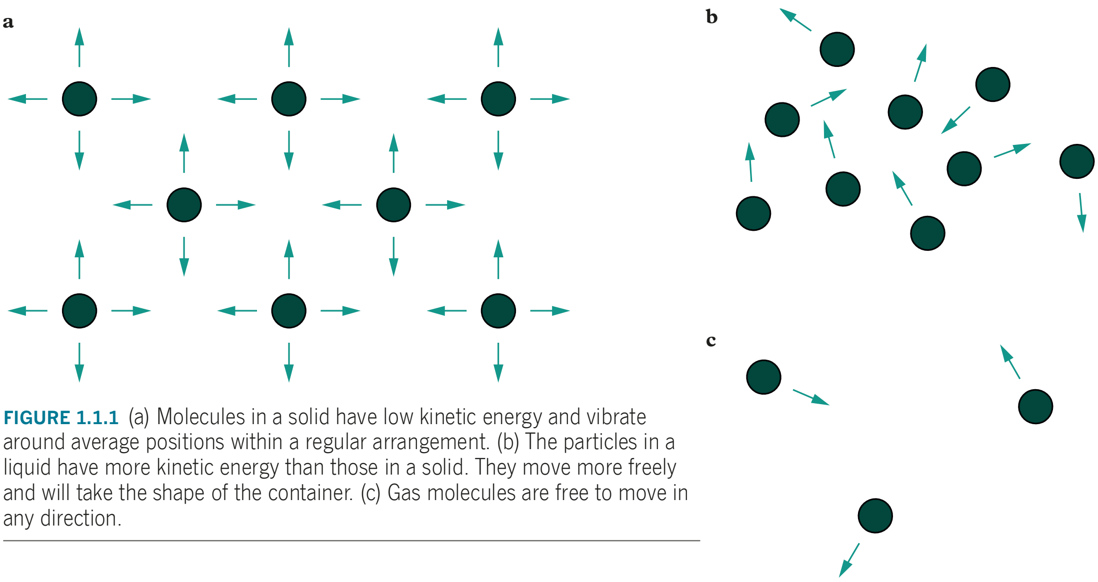


FIGURE 1.1.1 (a) Molecules in a solid have low kinetic energy and vibrate around average positions within a regular arrangement. (b) The particles in a liquid have more kinetic energy than those in a solid. They move more freely and will take the shape of the container. (c) Gas molecules are free to move in any direction.

PHYSICSFILE

States of matter: Plasma

The three phases (states) of matter that you normally come across are solid, liquid and gas. These are generally all that are discussed in secondary science.

There are in fact four phases of matter that are observable in everyday life—solid, liquid, gas, and plasma. Plasma exists when matter is heated to very high temperatures and electrons are freed (ionisation). A gas that is ionised and has an equal number of positive and negative charges is called plasma. The interior of stars consists of plasma. In fact, most of the matter in the universe is plasma (Figure 1.1.2).

The fifth known state of matter is known as a Bose–Einstein condensate. This is a state of matter that exists very close to absolute zero, a point at which molecular motion almost stops. Atoms begin to clump together and matter exhibits many of the properties of a super fluid; that is, it flows without friction.

FIGURE 1.1.2 99.9 per cent of the visible universe is made up of plasma.

THE KINETIC PARTICLE MODEL, INTERNAL ENERGY AND TEMPERATURE

The kinetic particle model can be used to explain the idea of heat as a transfer of energy. **Heat** (measured in joules) is the transfer of **thermal energy** from a hotter body to a colder one. Heating is observed by the change in **temperature**, the change of state or the expansion of a substance.

When a solid substance is 'heated', the particles within the material gain either **kinetic energy** (i.e. they move faster) or **potential energy** (the energy stored in the bonds between particles).

The term heat refers to energy that is being transferred (moved). So it is incorrect to talk about heat contained in a substance. The term **internal energy** refers to the total kinetic and potential energy of the particles within a substance. Heating (the transfer of thermal energy) changes the internal energy of a substance by affecting the kinetic energy and/or potential energy of the particles within the substance. The movement of the whole object due to kinetic energy is ordered: the object moves back and forth and its behaviour can be modelled. In comparison, the internal energy of a system is associated with the rapid and chaotic motion of the particles—it concerns the behaviour of a large number of particles that all have their own kinetic and potential energy.

- Heating is a process that always transfers thermal energy from a hotter substance to a colder substance.
- Heat is measured in joules (J).
- Temperature is related to the average kinetic energy of the particles in the substance. The faster the particles move, the higher the temperature of the substance.

PHYSICS IN ACTION

Energy

Energy is a very important concept in the study of the physical world, and is a focus in all areas of scientific study. Later chapters investigate energy in more detail.

Energy is a measure of an object's ability to do **work**. For example, raising an object's temperature or lifting an object is referred to as doing work. Work is measured in joules. The symbol for joules is J. Figure 1.1.3 shows the amount of joules available from some energy sources.

Kinetic energy is the energy of movement. It is equal to the amount of work needed to bring an object from rest to its present speed or to return it to rest. Potential energy is stored energy. There are many forms of potential energy, for example gravitational, nuclear, elastic and chemical. Chemical potential energy is associated with the bonds between the particles of a substance. An increase in the potential energy of particles in a substance results in movement of the particles from their equilibrium positions.

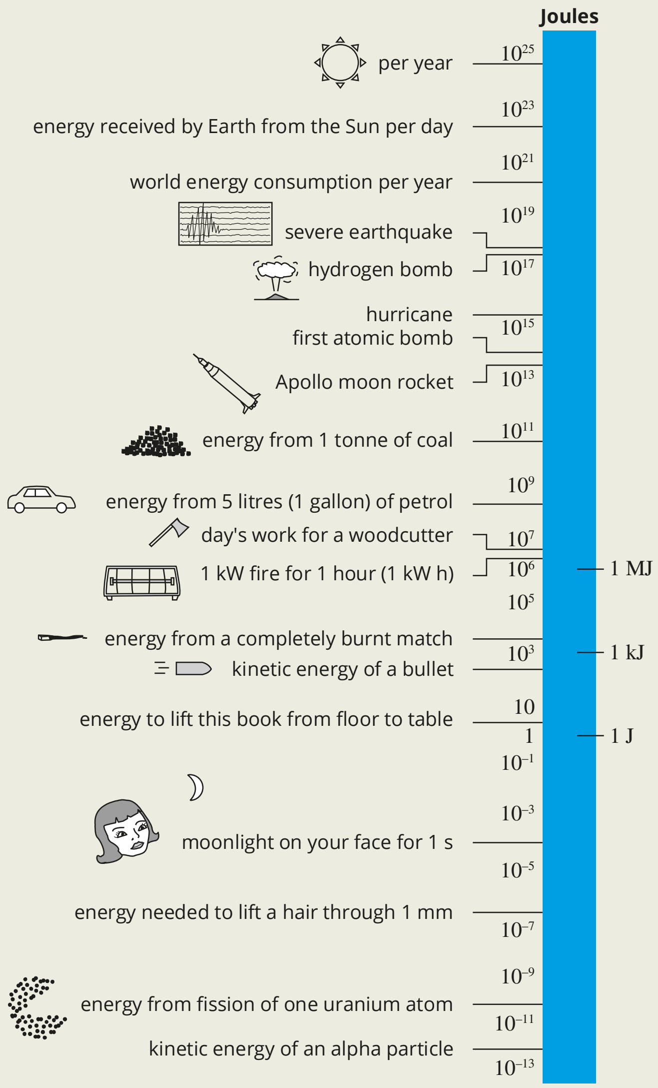


FIGURE 1.1.3 The comparative amounts of energy available from several sources.

Using the kinetic particle model, an increase in the total internal energy of the particles in a substance will result in an increase in temperature if there is a net gain in kinetic energy. Hot air balloons are an example of this process in action. The air in a hot air balloon is heated by a gas burner to a maximum of 120°C. The nitrogen (78%) and oxygen (21%) molecules in the hot air gain energy and so move a lot faster. The air in the balloon becomes less dense than the surrounding air, causing the balloon to float as seen in Figure 1.1.4.

Sometimes heating results only in the change of state or expansion of an object, and not a change in temperature. In these cases, the total internal energy of the particles has still increased but only the potential energy has increased, not the kinetic energy.

For instance, particles in a solid being heated will continue to be mostly held in place, due to the relatively strong interparticle forces. For the substance to change state from solid to liquid, it must receive enough energy to separate the particles from each other and disrupt the regular arrangement of the solid. During this 'phase change' process, the energy is used to overcome the strong interparticle forces, but not to change the overall speed of the particles. In this situation, the temperature does not change. This will be discussed in more detail in Section 1.3 'Latent heat'.

MEASURING TEMPERATURE

Only four centuries ago, there were no thermometers and people described heating effects by vague terms such as hot, cold and lukewarm. In about 1593, Italian inventor Galileo Galilei made one of the first thermometers. His 'thermoscope' was not particularly accurate as it did not take into account changes in air pressure, but it did suggest some basic principles for determining a suitable scale of measurement. His work used two fixed points: the hottest day of summer and the coldest day of winter. A scale like this is referred to as an arbitrary scale, because the fixed points are randomly chosen.

Celsius and Fahrenheit scales

Two of the better known arbitrary temperature scales are the Fahrenheit and Celsius scales. Gabriel Fahrenheit of Germany invented the first mercury thermometer in 1714. While Fahrenheit is still used in the United States of America to measure temperature, the system used in most countries of the world is now the Celsius scale. Two well-known fixed points for the Celsius scale are the standard freezing point of water, 0 degrees Celsius, and the boiling point of water, 100 degrees Celsius (at 1 atmosphere pressure).

Kelvin scale

Absolute scales are different from arbitrary scales. For a scale to be regarded as 'absolute', it should have no negative values. The fixed points must be reproducible and have zero as the lowest value. The kelvin scale for measuring temperature is an example of an absolute scale.

In developing the absolute temperature scale, the triple point of water provided one reliable fixed point. This is a point where the combination of temperature and air pressure allows all three states of water to coexist. For water, the triple point is only slightly above the standard freezing point at approximately 0.01°C and provides a unique and repeatable temperature with which to adjust the Celsius scale.

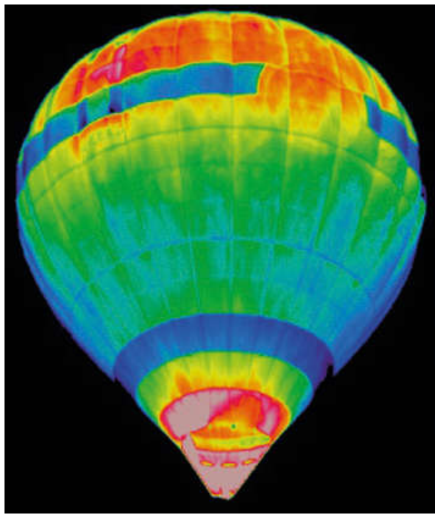


FIGURE 1.1.4 (a) Nitrogen and oxygen molecules gain energy when the air is heated, lowering the density of the air and causing the hot air balloon to rise off the ground. (b) A thermal image shows the temperature of the air inside the balloon, with the hotter areas showing as red.

PHYSICSFILE

Unit conventions in physics

The unit for energy, the joule, is named after James Joule in recognition of his work. When a unit is named after a person, its symbol is usually a capital letter but the unit name is always lower case, e.g. joule (J) is named after James Joule, newton (N) is named after Isaac Newton, and kelvin (K) is named after Lord Kelvin.

Exceptions are degrees Celsius (°C) and degrees Fahrenheit (°F), which also include a degree symbol and have a capital letter for the unit.

Units not named after people usually have both the symbol and the name in lowercase, e.g. metre (m), second (s).

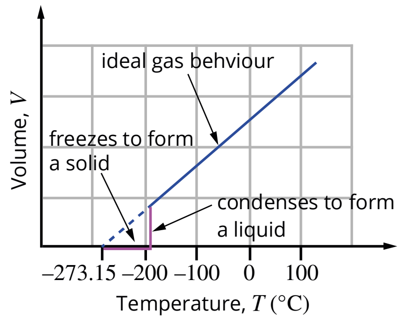


FIGURE 1.1.6 Gases have smaller volumes as they cool. This relationship is linear.

PHYSICSFILE

Close to absolute zero

As temperatures get close to absolute zero, atoms start to behave in weird ways. Since the French physicist Guillaume Amontons first proposed the idea of an absolute lowest temperature in 1699, physicists have theorised about the effects of such a temperature and how it could be achieved. The laws of physics dictate that absolute zero itself can be approached but not reached. In 2003, researchers from NASA and MIT, in the United States of America, succeeded in cooling sodium atoms to one billionth of a degree above absolute zero. At this temperature, all elementary particles merge into a single state (a Bose–Einstein condensate), losing their separate properties and behaving as a single 'super atom', a state first proposed by Einstein 70 years earlier.

Absolute zero

Experiments indicate that there is a limit to how cold things can get. The kinetic theory suggests that if a given quantity of gas is cooled, its volume decreases. The volume can be plotted against temperature and results in a straight-line graph as shown in Figure 1.1.6. Extrapolating (extending) the line to where the volume is zero gives a theoretical value of absolute zero.

- Absolute zero = 0 K = –273.15°C
- All molecular motion ceases at absolute zero. This is the coldest temperature possible.

The absolute or **kelvin** temperature scale is based on absolute zero and the triple point of water. See Figure 1.1.5 for a comparison of the kelvin and Celsius scales.

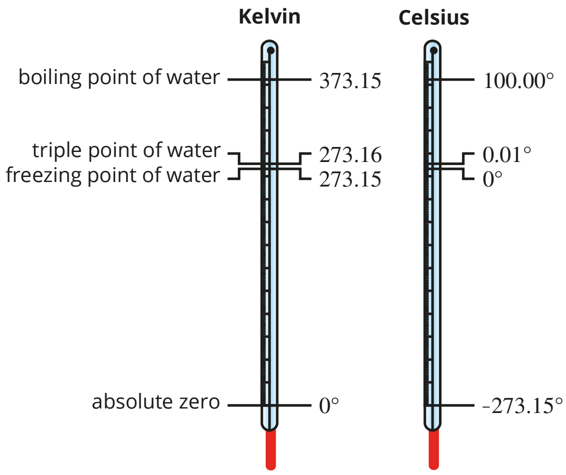


FIGURE 1.1.5 Comparison of the kelvin and Celsius scales. Note that there are no negative values on the kelvin scale.

- The freezing point of water (0°C) is equivalent to 273.15 K (kelvin). This is often approximated to 273 K.
- The size of each unit, 1°C or 1 K, is the same.
- The word 'degree' and the degree symbol are not used with the kelvin scale.
- To convert a temperature from degrees Celsius to kelvin: add 273.
- To convert a temperature from kelvin to degrees Celsius: subtract 273.
- 0 degrees Celsius (0°C) is the freezing point of water at standard atmospheric pressure.
- 100°C is the boiling point of water at standard atmospheric pressure.

THERMAL EQUILIBRIUM AND ENERGY TRANSFER

Whenever two materials, initially at different temperatures, come into contact (or are mixed) then thermal energy will transfer from the hotter material to the colder one until both are the same temperature. For example, if a piece of frozen fruit is placed in a container of warm water, energy is transferred from the water to the fruit. The fruit gains energy and warms up. The water loses energy and cools down. Eventually the transfer of energy between the fruit and water will stop. This point is called **thermal equilibrium,** and the fruit and water will be at the same temperature. Two objects in thermal equilibrium with each other must be at the same temperature. This is referred to as the zeroth law of thermodynamics.

In this example, it's likely that the fruit will no longer be frozen. Moving to a temperature where thermal equilibrium occurs has involved not only a change in temperature but also a change of state. This is true of any thermal energy transfer where systems or objects in contact reach the same temperature. The energy transfer may involve changes of state as well as a change in temperature. Assuming no energy is lost from the systems, the total thermal energy will remain the same.

MECHANICAL WORK AND THERMAL ENERGY

The first law of thermodynamics states that energy simply changes from one form to another and the total energy in an isolated system remains constant. The internal energy of the system can be changed by heating or cooling (transfer from or to another system that's in contact), or by doing mechanical work on or by the system.

The internal energy (U) of a system is defined as the total internal kinetic and potential energy of the system. As the average kinetic energy of a system is related to its temperature and the potential energy of the system is related to the state, then a change in the internal energy of a system means that either the temperature changes or the state changes.

If heat (Q) is added to the system, then the internal energy (U) rises by either increasing temperature or changing state from solid to liquid or liquid to gas. Similarly, if mechanical work (W) is done on a system, then the internal energy rises and the system will once again increase in temperature or will change state by melting or boiling. When heat is added to a system or work is done on a system, ΔU is positive.

If heat (Q) is removed from the system, then the internal energy (U) decreases by either decreasing temperature or changing state from liquid to solid or gas to liquid. Similarly, if mechanical work (W) is done by the system, then the internal energy decreases and the system will decrease in temperature or change state by condensing or solidifying. When heat is removed from a system or work is done by a system, ΔU is negative.

If heat is added to the system and work is done by the system, then whether the internal energy increases or decreases will depend on the magnitude of the energy into the system compared to the magnitude of the energy out of the system.

Worked example 1.1.1

CALCULATING THE CHANGE IN INTERNAL ENERGY

A 1 L beaker of water has 25 kJ of work done on it and loses 30 kJ of thermal energy to the surroundings. What is the change in energy of the water?

Thinking	Working
Heat is removed from the system, so Q is negative. Work is done on the system, so W is positive.	$\Delta U = Q + W$ $= (-30) + (+25)$
Note that the units are kJ, so express the final answer in kJ.	$\Delta U = -5$ kJ

Worked example: Try yourself 1.1.1

CALCULATING THE CHANGE IN INTERNAL ENERGY

A student places a heating element and a paddle-wheel apparatus in an insulated container of water. She calculates that the heating element transfers 2530 J of thermal energy to the water and the paddle does 240 J of work on the water. Calculate the change in internal energy of the water.

PHYSICSFILE

The zeroth law of thermodynamics

The topic of thermal physics involves the phenomena associated with the energy transfer between objects at different temperatures. Since the nineteenth century, scientists have developed four laws for this subject. The first, second and third laws were known and understood for some time. Then another law was determined. This final law was considered to be so important that it was decided to place it before the first, and so it is called the zeroth (0^{th}) law of thermodynamics.

Any change in the internal energy (ΔU) of a system is equal to the energy added by heating ($+Q$) or removed by cooling ($-Q$), minus the work done on ($+W$) or by ($-W$) the system.

$\Delta U = Q + W$

PHYSICSFILE

Mechanical work

Physics has a different definition of work to that in common speech—it's not directly related to things like 'going to work' or 'homework'. Mechanical work is a measure of the energy transferred by a force. While heat transfer involves energy transfer, it is not considered to be mechanical work, since there is no measurable force involved in the transfer. Note that heat, internal energy and work all have the same unit of joules (J).

The ideas behind mechanical work are further developed in Chapter 9.

1.1 Review

SUMMARY

- The kinetic particle theory proposes that all matter is made of atoms or molecules (particles) that are in constant motion.
- In solids, the attractive and repulsive forces hold the particles in fixed positions, usually in a regular arrangement or lattice. These particles are not completely still—they vibrate about average positions.
- In liquids, there is still a balance of attractive and repulsive forces between particles but the particles have more freedom to move around. Liquids maintain a fixed volume.
- In gases, the particle speeds are high enough that, when particles collide, the attractive forces are not strong enough to keep them close together. The repulsive forces cause the particles to move off in other directions.
- Internal energy refers to the total kinetic and potential energy of the particles within a substance.
- Temperature is related to the average kinetic energy of the particles in a substance.
- Heating is a process that always transfers thermal energy from a hotter substance to a colder substance.
- Temperatures can be measured in degrees Celsius (°C) or kelvin (K).
- Absolute zero is called simply 'zero kelvin' (0 K) and it is equal to –273.15°C.
- The size of each increment of temperature is the same—an increase of 1°C is equal to an increase of 1 K.
- To convert from Celsius to kelvin: add 273; to convert from kelvin to Celsius: subtract 273 (to 3 significant figures).
- Whenever two materials initially at different temperatures come into contact (or are mixed), thermal energy will transfer from the hotter material to the colder one until both are the same temperature. This is referred to as thermal equilibrium
- Any change in the internal energy (ΔU) of a system is equal to the energy added by heating ($+Q$) or removed by cooling ($-Q$), plus the work done on ($+W$) or by ($-W$) the system: $\Delta U = Q + W$.

KEY QUESTIONS

1 Which of the following is true of a solid?
 A Particles are moving around freely.
 B Particles are not moving.
 C Particles are vibrating in constant motion.
 D A solid is not made up of particles.

2 An uncooked chicken is placed into an oven that has been preheated to 180°C. Order the following statements to describe what happens in the first few seconds the chicken is placed in the oven.
 - The chicken and the air in the oven are in thermal equilibrium.
 - Thermal energy flows from the hot air into the chicken.
 - The chicken and the air in the oven are not in thermal equilibrium.

3 Which of the following temperature(s) cannot possibly exist? (More than one answer is possible.)
 A 1 000 000°C **B** –50°C **C** –50 K **D** –300°C

4 A tank of pure helium is cooled to its freezing point of –272.2°C. Describe the energy of the helium particles at this temperature.

5 Covert the following temperatures:
 a 30°C into kelvin
 b 375 K into degrees Celsius.

6 Tank A is filled with hydrogen gas at 0°C and another tank, B, is filled with hydrogen gas at 300 K. Describe the difference in the average kinetic energy of the hydrogen particles in each tank.

7 Sort the following temperatures from coldest to hottest:
 - freezing point of water
 - 100 K
 - absolute zero
 - –180 °C
 - 10 K

8 A hot block of iron does 50 kJ of work on a cold floor. The block of iron also transfers 20 kJ of heat energy to the air. Calculate the change in internal energy (ΔU) of the iron block in kJ.

9 A chef vigorously stirs a pot of cold water and does 150 J of work on the water. The water also gains 75 J of thermal energy from the surroundings. Calculate the change in energy of the water.

10 A scientist very carefully does mechanical work on a container of liquid sodium. The liquid sodium loses 300 J of energy to its surroundings but gains 250 J of energy overall. How much work did the scientist do on the liquid sodium?

1.2 Specific heat capacity

A small amount of water in a kettle will experience a greater change in temperature than a larger volume if heated for the same time. A metal object left in the sunshine gets hotter faster than a wooden object. Large heaters warm rooms faster than small ones.

These simple observations suggest that the mass, material, and the amount of energy transferred influence any change of temperature.

CHANGING TEMPERATURE

The temperature of a substance is a measure of the average kinetic energy of the particles inside the substance. To increase the temperature of the substance, the kinetic energy of its particles must increase. This happens when heat is transferred to that substance. The amount the temperature increases depends on several factors.

The greater the mass of a substance, the greater the energy required to change the kinetic energy of all the particles. So, the heat required to raise the temperature by a given amount is proportional to the mass of the substance.

$\Delta Q \propto m$

where ΔQ is the heat energy transferred in joules (J)

m is the mass of material being heated in kilograms (kg).

The more heat that is transferred to a substance, the more the temperature of that substance increases. The amount of energy transferred is therefore proportional to the change in temperature.

$\Delta Q \propto \Delta T$

where ΔT is the change in temperature in °C or K.

Heating experiments using different materials will confirm that these relationships hold true regardless of the material being heated. However, heating the same masses of different materials will show that the amount of energy required to heat a given mass of a material through a temperature change also depends on the nature of the material being heated. For example, a volume of water requires more energy to change its temperature by a given amount compared with the same volume of methylated spirits. For some materials, temperature change occurs more easily than for others.

The **specific heat capacity** of a material, c, is the amount of energy that must be transferred to change the temperature of 1 kg of the material by 1°C or 1 K.

Combining these observations, the amount of energy added to or removed from the substance is proportional to the change in its temperature, its mass and its specific heat capacity (provided a material does not change state). The specific heat capacity of a material changes when the material changes state.

As an equation:

$Q = mc\Delta T$

where Q is the heat energy transferred in joules (J)

m is the mass in kilograms (kg)

ΔT is the change in temperature in °C or K

c is the specific heat capacity of the material (J kg^{-1} K^{-1}).

Table 1.2.1 lists the specific heat capacities for some common materials. It also includes the average value for the human body, taking into account the various materials within the body and the percentage that each material contributes to the body's total mass.

TABLE 1.2.1 Approximate specific heat capacities of common substances.

Material	c ($J\,kg^{-1}\,K^{-1}$)
human body	3500
methylated spirits	2500
air	1000
aluminium	900
glass	840
iron	440
copper	390
brass	370
lead	130
mercury	140
ice (water)	2100
liquid water	4180
steam (water)	2000

Worked example 1.2.1

CALCULATIONS USING SPECIFIC HEAT CAPACITY

A hot water tank contains 135 L of water. Initially the water is at 20.0°C. Calculate the amount of energy that must be transferred to the water to raise the temperature to 70.0°C.

Thinking	Working
Calculate the mass of water. 1 L of water = 1 kg	Volume = 135 L So mass of water = 135 kg
ΔT = final temperature – initial temperature	$\Delta T = 70.0 - 20.0 = 50.0$°C
From Table 1.2.1, $c_{water} = 4180\,J\,kg^{-1}\,K^{-1}$. Use the equation $Q = mc\Delta T$.	$Q = mc\Delta T$ $= 135 \times 4180 \times 50.0$ $= 28\,215\,000$ J $= 2.82 \times 10^7$ J

Worked example: Try yourself 1.2.1

CALCULATIONS USING SPECIFIC HEAT CAPACITY

A bath contains 75 L of water. Initially the water is at 50°C. Calculate the amount of energy that must be transferred from the water to cool the bath to 30°C.

PHYSICSFILE

The mass of water

Since water is a familiar material, many of the examples in this section use it as the liquid being heated. One kilogram of pure water has a volume of 1 litre at 4°C.

Worked example 1.2.2

COMPARING SPECIFIC HEAT CAPACITIES

Different states of matter of the same substance have different specific heat capacities. What is the ratio of the specific heat capacity of liquid water to that of ice?	
Thinking	**Working**
Table 1.2.1 has the specific heat capacities of water in different states.	$c_{water} = 4180\,\text{J kg}^{-1}\,\text{K}^{-1}$ $c_{ice} = 2100\,\text{J kg}^{-1}\,\text{K}^{-1}$
Divide the specific heat of water by the specific heat of ice.	$\text{Ratio} = \dfrac{c_{water}}{c_{ice}}$ $= \dfrac{4180}{2100}$
Note that ratios have no units since the unit of each quantity is the same and cancels out.	Ratio ≈ 2

Worked example: Try yourself 1.2.2

COMPARING SPECIFIC HEAT CAPACITIES

What is the ratio of the specific heat capacity of liquid water to that of steam?

PHYSICSFILE

Specific heat capacity of water

One of the notable values in the table of specific heat capacities is the high value for water. It is 10 times, or an order of magnitude, higher than those of most metals listed. The specific heat capacity of water is higher than those of most common materials. As a result, water makes a very useful cooling and heat storage agent, and is used in areas such as generator cooling towers and car-engine radiators.

Life on Earth also depends on the specific heat capacity of water. About 70% of the Earth's surface is covered by water, and these water bodies can absorb large quantities of thermal energy without great changes in temperature. Oceans both heat up and cool down more slowly than the land areas next to them. This helps to maintain a relatively stable range of temperatures for life on Earth.

Scientists are now monitoring the temperatures of the deep oceans to determine how the ability of oceans to store large amounts of energy may affect climate change.

1.2 Review

SUMMARY

- When heat is transferred to or from a system or object, the temperature change depends upon the amount of energy transferred, the mass of the material(s) and the specific heat capacity of the material(s): $Q = mc\Delta T$

  where Q is the heat energy transferred in joules (J)

  m is the mass of material being heated in kilograms (kg)

  ΔT is the change in temperature (°C or K)

  c is the specific heat capacity of the material ($\text{J kg}^{-1}\text{ K}^{-1}$).
- A substance will have different specific heat capacities at different states (solid, liquid, gas).

KEY QUESTIONS

1 Equal masses of water and aluminium are heated through the same temperature range. Using the values of *c* from Table 1.2.1 on page 12, which material requires the most energy to achieve this result?

2 Which has the most thermal energy: 10 kg of iron at 20°C or 10 kg of aluminium at 20°C?

3 100 mL of water is heated to change its temperature from 15°C to 20°C. How much energy is transferred to the water to achieve this temperature change?

4 150 mL of water is heated from 10°C to 50°C. What amount of energy is required for this temperature change to occur?

5 For a 1 kg block of aluminium, *x* J of energy are needed to raise the temperature by 10°C. How much energy, in J, is needed to raise the temperature by 20°C?

6 Equal masses of aluminium and water absorb equal amounts of energy. What is the ratio of the temperature rise of the aluminium to that of water?

7 Which one or more of the following statements about specific heat capacity is true?

A All materials have the same specific heat capacity when in solid form.

B The specific heat capacity of a liquid form of a material is different from that of the solid and gas forms.

C Good conductors of heat generally have high specific heat capacities.

D Specific heat capacity is independent of temperature.

8 If 4.0 kJ of energy is required to raise the temperature of 1.0 kg of paraffin by 2.0°C, how much energy (in kJ) is required to raise the temperature of 5.0 kg of paraffin by 1.0°C?

9 A cup holds 250 mL of water at 20°C. 10.5 kJ of heat energy is transferred to the water. What temperature does the water reach after the heat is transferred?

10 A block of iron is left to cool. After cooling for a short time, 13.2 kJ of energy has been transferred away from the block of iron and its temperature has decreased by 30°C. What is the mass of the block of iron?

1.3 Latent heat

If water is heated, its temperature will rise. If enough energy is transferred to the water, eventually the water will boil. The water changes state (from liquid to gas). The **latent heat** is the energy released or absorbed during a change of state. Latent means hidden or unseen. While a substance changes state, its temperature remains constant. The energy used in, say, melting ice into water is hidden in the sense that the temperature doesn't rise while the change of state is occurring.

ENERGY AND CHANGE OF STATE

Figure 1.3.1 shows a heating curve for water, illustrating how the temperature of water changes as energy is added at a constant rate. Although the rate at which the energy is added is constant, the increase in temperature is not always constant. There are sections of increasing temperature, and sections where the temperature remains unchanged (the horizontal sections) while the material changes state. The temperature of the water remains constant during the change in state from ice to liquid water and again from liquid water to steam.

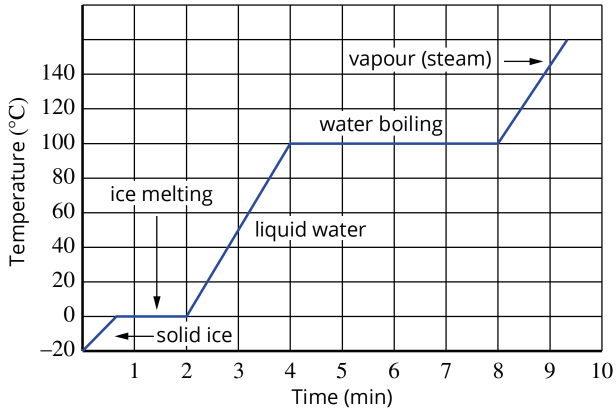


FIGURE 1.3.1 A heating curve for water.

Latent heat

The energy needed to change the state of a substance (e.g. solid to liquid, liquid to gas) is called latent heat. Latent heat is the 'hidden' energy that must be added or removed from a material in order for the material to change state.

The latent heat is calculated using the equation:

$Q = mL$

where Q is the heat energy transferred in joules (J)

m is the mass in kilograms (kg)

L is the latent heat (J kg^{-1}).

LATENT HEAT OF FUSION (MELTING)

As thermal energy is transferred to a solid, the temperature of the solid increases. The particles within the solid gain internal energy (as kinetic energy and some potential energy) and their speed of vibration increases. At the point where the solid begins to melt, the particles move further apart, reducing the strength of the bonds holding them in place. At this point, instead of increasing the temperature, the extra energy increases the potential energy of the particles, reducing the interparticle or intermolecular forces. No change in temperature occurs as all the extra energy supplied is used in reducing these forces between particles.

The amount of energy required to melt a solid is the same as the amount of potential energy released when the liquid re-forms into a solid. It is termed the **latent heat of fusion**.

The amount of energy required will depend on the particular solid.

> For a given mass of a substance:
>
> heat energy transferred = mass of substance × specific latent heat of fusion
>
> $Q = mL_{fusion}$
>
> where Q is the heat energy transferred in joules (J)
>
> m is the mass in kilograms (kg)
>
> L_{fusion} is the latent heat of fusion in J kg^{-1}.

It takes almost 80 times as much energy to turn 1 kg of ice into water (with no temperature change) as it does to raise the temperature of 1 kg of water by 1°C. It takes a lot more energy to overcome the large intermolecular forces within the ice than it does to simply add kinetic energy in raising the temperature.

The latent heats of fusion for some common materials are shown in Table 1.3.1.

TABLE 1.3.1 The latent heats of fusion for some common materials.

Substance	Melting point (°C)	L_{fusion} (J kg^{-1})
water	0	3.34×10^5
oxygen	–219	0.14×10^5
lead	327	0.25×10^5
ethanol	–114	1.05×10^5
silver	961	0.88×10^5

Worked example 1.3.1

LATENT HEAT OF FUSION

How much energy must be removed from 2.50 L of water at 0.00°C to produce a block of ice at 0.00°C? Express your answer in kJ.

Thinking	Working
Cooling from liquid to solid involves the latent heat of fusion, where the energy is removed from the water. Calculate the mass of water involved.	1 L of water = 1 kg, so 2.50 L = 2.50 kg
Use Table 1.3.1 to find the latent heat of fusion for water.	$L_{fusion} = 3.34 \times 10^5$ J kg^{-1}
Use the equation $Q = mL_{fusion}$.	$Q = mL_{fusion}$ $= 2.50 \times 3.34 \times 10^5$ $= 8.35 \times 10^5$ J
Convert to kJ.	$Q = 8.35 \times 10^2$ kJ

Worked example: Try yourself 1.3.1

LATENT HEAT OF FUSION

How much energy must be removed from 5.5 kg of liquid lead at 327°C to produce a block of solid lead at 327°C? Express your answer in kJ.

LATENT HEAT OF VAPORISATION (BOILING)

It takes much more energy to convert a liquid to a gas than it does to convert a solid to a liquid. This is because, to convert a liquid to a gas, the intermolecular bonds must be broken. During the change of state, the energy supplied is used solely in overcoming intermolecular bonds. The temperature will not rise until all the material in the liquid state is converted to a gas, assuming that the liquid is evenly heated. For example, when liquid water is heated to boiling point, a large amount of energy is required to change its state from liquid to steam (gas). The temperature will remain at 100°C until all the water has turned into steam. Once the water is completely converted to steam, then the temperature can start to rise again.

The amount of energy required to change a liquid to a gas is exactly the same as the potential energy released when the gas returns to a liquid. It is called the **latent heat of vaporisation**.

The amount of energy required will depend on the particular substance.

For a given mass of a substance:

heat energy transferred = mass of substance × latent heat of vaporisation

$Q = mL_{vapour}$

where Q is the heat energy transferred in joules (J)

m is the mass in kilograms (kg)

L_{vapour} is the latent heat of vaporisation ($J\,kg^{-1}$).

Note that, in just about every case, the latent heat of vaporisation of a substance will be different to the latent heat of fusion for that substance. Some latent heat of vaporisation values are listed in Table 1.3.2.

In many instances, it is necessary to consider the energy required to heat a substance and to also change its state. Problems like this are solved by considering the rise in temperature separately from the change of state.

TABLE 1.3.2 The latent heat of vaporisation of some common materials.

Substance	Boiling point (°C)	L_{vapour} ($J\,kg^{-1}$)
water	100	22.5×10^5
oxygen	−183	2.2×10^5
lead	1750	9.0×10^5
ethanol	78	8.7×10^5
silver	2193	23.0×10^5

PHYSICSFILE

Extinguishing fire

The latent heat of vaporisation of water is very high. This is due to the molecular structure of the water. This characteristic of water makes it very useful for extinguishing fires. That's because water can absorb vast amounts of thermal energy before it evaporates. By pouring water onto a fire, energy is transferred away from the fire to heat the water. Then, even more (in fact much more) heat is transferred away from the fire as the liquid water is converted into steam.

The rate of evaporation of a liquid depends on:

- the volatility of the liquid: more volatile liquids evaporate faster
- the surface area: greater evaporation occurs when greater surface areas are exposed to the air
- the temperature: hotter liquids evaporate faster
- the humidity: less evaporation occurs in more humid conditions
- air movement: if a breeze is blowing over the liquid's surface, evaporation is more rapid.

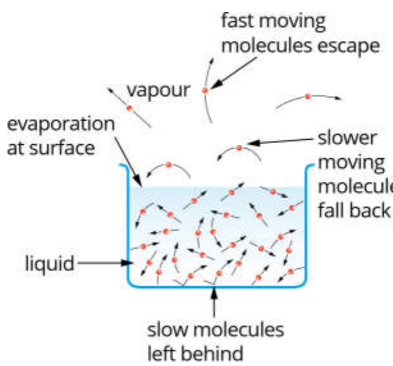


FIGURE 1.3.2 Fast-moving molecules with high kinetic energy can escape the liquid, leaving molecules with lower kinetic energy behind.

Worked example 1.3.2

CHANGE IN TEMPERATURE AND STATE

50.0 mL of water is heated from a room temperature of 20.0°C to its boiling point at 100.0°C. It is boiled at this temperature until it is completely evaporated. How much energy in total was required to raise the temperature and boil the water?

Thinking	Working
Calculate the mass of water involved.	50.0 mL of water = 0.0500 kg
Find the specific heat capacity of water from Table 1.2.1.	$c = 4180\ \text{J kg}^{-1}\ \text{K}^{-1}$
Use the equation $Q = mc\Delta T$ to calculate the heat energy required to change the temperature of water from 20.0°C to 100.0°C.	$Q = mc\Delta T$ $= 0.0500 \times 4180 \times (100.0 - 20.0)$ $= 16720$ $= 1.67 \times 10^4\ \text{J}$
Find the specific latent heat of vaporisation of water.	$L_{vapour} = 22.5 \times 10^5\ \text{J kg}^{-1}$
Use the equation $Q = mL_{vapour}$ to calculate the latent heat required to boil water.	$Q = mL_{vapour}$ $= 0.0500 \times 22.5 \times 10^5$ $= 112500$ $= 1.13 \times 10^5\ \text{J}$
Find the total energy required to raise the temperature and change the state of the water.	Total $Q = (1.67 \times 10^4) + (1.13 \times 10^5)$ $= 1.29 \times 10^5\ \text{J}$

Worked example: Try yourself 1.3.2

CHANGE IN TEMPERATURE AND STATE

3 L of water is heated from a fridge temperature of 4°C to its boiling point at 100°C. It is boiled at this temperature until it is completely evaporated. How much energy in total is required to raise the temperature and boil the water?

EVAPORATION AND COOLING

If you spill some water on the floor then come back in a couple of hours, the water will probably be gone. It will have evaporated. It has changed from a liquid into a vapour at room temperature in a process called **evaporation**. The reason for this is that the water particles, if they have sufficient energy, can escape through the surface of the liquid into the air. Over time, no liquid remains.

Evaporation is more noticeable in **volatile** liquids such as methylated spirits, mineral turpentine, perfume and liquid paper. The surface bonds are weaker in these liquids and they evaporate rapidly. This is why you should never leave the lid off bottles of these liquids. They are often stored in narrow-necked bottles for this reason.

Whenever evaporation occurs, higher-energy particles escape the surface of the liquid, leaving the lower-energy particles behind, as is shown in Figure 1.3.2. As a result, the average kinetic energy of the particles remaining in the liquid drops and the temperature decreases. Humans use this cooling principle when sweating to stay cool. When rubbing alcohol is dabbed on your arm before an injection, the cooling of the volatile liquid numbs your skin.

PHYSICS IN ACTION

The development of refrigeration

Most people are familiar with the cooling effect of getting out of a pool and into a breeze. It is due to evaporation and can make you feel cold even on a warm, sunny day. Figure 1.3.2 illustrates what happens at the surface of a liquid during evaporation.

A molecule at the surface of the liquid gains energy, as thermal energy, from the surroundings. This energy is transformed to kinetic energy of the molecule. The extra kinetic energy allows the molecule to overcome the surface tension of the liquid and change state to a gas. If the molecule still has enough energy, it will stay in the gas state and escape the liquid entirely, 'evaporating'. As only those molecules with the highest kinetic energy escape from the liquid, the average kinetic energy of the molecules remaining in the liquid decreases. A lower average kinetic energy means a lower temperature.

Early refrigerators used the evaporative cooling principle to stop foods such as meat and fruit from spoiling. The Coolgardie meat safe (Figure 1.3.3) was invented in the small Western Australian mining town near Kalgoorlie, after which it was named.

FIGURE 1.3.3 Early model of the Coolgardie meat safe.

The Coolgardie meat safe was a simple cupboard made of wire mesh and a wooden or pressed-metal frame. A water tray was placed on top of the cupboard and the top of a hessian bag was attached to the water tray to soak up water. The wet hessian bag hung down, covering the safe like a flap. The Coolgardie meat safe was usually placed on a veranda or similar place where there was a breeze. The breeze passed through the wet hessian and the water evaporated.

The energy required for evaporation to occur (latent heat) was drawn from the passing breeze or from the metal and hence the air or metal was cooled. This would cool the safe and thus the food inside it.

Modern refrigerators and air conditioners still rely on evaporation to cool food or a room, although the way it is used is quite different.

Refrigerators and air conditioners make things cold by removing energy from them. Energy is pumped from the space being cooled to the outside air. As a result, modern refrigeration systems can be called 'heat pumps'. The cooling process is shown in Figure 1.3.4.

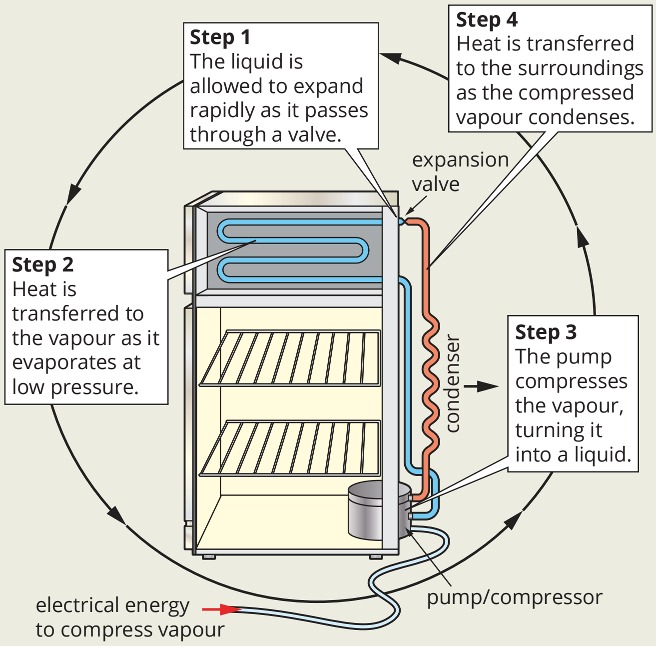


FIGURE 1.3.4 A refrigerator acts as a heat pump to remove energy from inside the refrigerator.

Inside a heat pump, a volatile liquid known as a refrigerant is circulated around a closed system of pipes by a pump. The pressure inside the evaporator pipes is reduced by an expansion valve (step 1), causing the refrigerant to evaporate. Energy is needed for this change of state to take place, i.e. the latent heat of vaporisation, so the system will absorb energy (step 2). As the gas evaporates, it absorbs energy from the surrounding air, making the air cooler.

In the condenser pipes, the process is reversed. The refrigerant gas is compressed and condenses to a liquid again (step 3). The change of state from a gas to liquid releases energy. This energy is released outside the refrigerator and heats the surrounding air (step 4). Try putting your hand in the space behind your fridge at home—what do you notice?

Modern reverse-cycle air conditioners and heat-pump water heaters reverse this process to heat homes and water. Heat energy is picked up from outside the house (or from a power source) and released inside when the vapour condenses.

1.3 Review

SUMMARY

- When a solid material changes state, energy is needed to separate the particles by overcoming the attractive forces between the particles.
- Latent heat is the energy required to change the state of 1 kg of material at a constant temperature.
- In general, for any mass of material the energy required (or released) is $Q = mL$

  where Q is the energy transferred in joules (J)

  m is the mass in kilograms (kg)

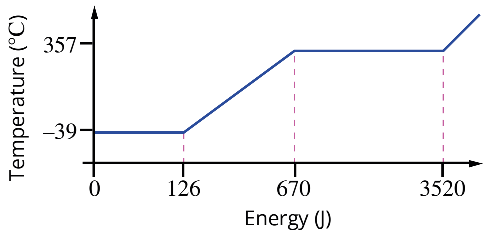
  L is the latent heat (J kg^{-1}).
- The latent heat of fusion, L_{fusion}, is the energy required to change 1 kg of a material between the solid and liquid states.
- The latent heat of vaporisation, L_{vapour}, is the energy required to change 1 kg of a material between the liquid and gaseous states.
- The latent heat of fusion of a material will be different to (and usually less than) the latent heat of vaporisation for that material.
- Evaporation is when a liquid turns into gas at room temperature. The temperature of the liquid falls as this occurs.
- The rate of evaporation depends on the volatility, temperature and surface area of the liquid, and the presence of a breeze.

KEY QUESTIONS

Refer to the values in Tables 1.3.1 and 1.3.2 on pages 16 and 17. You may also need to refer to Table 1.2.1 on page 12.

The following information applies to questions 1–5.

The graph below represents the heating curve for mercury, a metal that is a liquid at normal room temperature. Thermal energy is added to 10 g of solid mercury, initially at a temperature of –39°C, until all of the mercury has evaporated.



1 Why does the temperature remain constant during the first part of the graph?

2 What is the melting point of mercury, in degrees Celsius?

3 What is the boiling point of mercury, in degrees Celsius?

4 From the graph, what is the latent heat of fusion of mercury?

5 From the graph, what is the latent heat of vaporisation of mercury?

6 How much heat energy must be transferred away from 100 g of steam at 100°C to change it completely to a liquid?

7 How many kJ of energy are required to melt exactly 100 g of ice initially at –4.00°C? Assume no loss of energy to surroundings.

8 Explain why hot water in a spa-pool evaporates more rapidly than cold water.

9 A painter spills some mineral turpentine onto a concrete floor. After a minute, most of the liquid is gone and the floor is cooler where the liquid was. Explain these observations.

1.4 Heating and cooling

Have you ever got into a bath only to find the water had gone cold? You probably then added hot water and waited, shivering, for the bath to get warm. In this section you will learn what happens to the energy from the hot water when it is added to the cold water, and how to calculate much hot water you would need to add to get to the right temperature.

As was explained in Section 1.1, when objects at different temperatures come into contact, thermal energy is transferred from the hotter object to the cooler object. Once the two objects reach the same temperature, they are said to be in thermal equilibrium.

This section focuses on solving heating problems involving the direct transfer of heat, and mixtures to reach thermal equilibrium.

THERMAL EQUILIBRIUM

Whenever two materials initially at different temperatures come into contact (or are mixed), thermal energy will transfer from the hotter material to the colder one until both are the same temperature. Both will end up at the same average temperature and the mixture will have reached thermal equilibrium. The final temperature will be somewhere between the two original temperatures and will depend on the relative mass of each material and their relative specific heat capacities. The final temperature will only be exactly halfway between the two original temperatures if the masses and the respective specific heat capacities of both materials are equal.

Assuming no loss or gain from the surrounding environment, the total energy would remain the same. This is an example of conservation of energy and is an essential concept when solving practical problems involving a transfer of energy.

Worked example 1.4.1

CALCULATING THERMAL EQUILIBRIUM

10.0 kg of water initially at 80.0°C is mixed with 30.0 kg of water initially at 20.0°C. What is the final temperature of the water once thermal equilibrium is reached?	
Thinking	**Working**
Total energy lost by hot water = total energy gained by cold water That is, the energy change, ΔQ, is equal for the hot and cold water. Use $\Delta Q = mc\Delta T$. Assume no loss to the surrounding environment.	$\Delta Q_{hot} = \Delta Q_{cold}$ $m_{hot}\ c\Delta T_{hot} = m_{cold}\ c\Delta T_{cold}$
Since specific heat capacity of the water will be the same on both sides of the equation, the equation can be simplified.	$m_{hot}\ \Delta T_{hot} = m_{cold}\ \Delta T_{cold}$
Substitute the known values and simplify for the equilibrium temperature, T.	$10.0 \times (80.0 - T) = 30.0 \times (T - 20.0)$ $800 - 10.0T = 30.0T - 600$ $800 + 600 = 30.0T + 10.0T$ $1400 = 40.0T$ $T = \frac{1400}{40.0} = 35.0°C$
Do a quick intuitive check. Does the answer make sense?	As most of water was colder, the final temperature should be closer to the temperature of the original colder water than to the temperature of the original hotter water.

Worked example: Try yourself 1.4.1

CALCULATING THERMAL EQUILIBRIUM

4.00 kg of water initially at 85.0°C is mixed with 3.00 kg of water initially at 25.0°C. What is the final temperature of the water once thermal equilibrium is reached?

Worked example 1.4.2

CALCULATING THERMAL EQUILIBRIUM

A 50.0 g piece of iron is heated over a flame for several minutes. The iron is then plunged into an insulated, closed container containing 1.00 L of cool water, originally at 15.0°C. When thermal equilibrium is reached, the temperature of the water is found to be 17.0°C. If no water changes state to become steam and there are no other energy losses, then what was the temperature of the iron just before it was immersed in the water?

Thinking	Working
Convert all masses to standard units (kg).	Mass of iron = 50.0 g = 0.0500 kg Mass of water = 1.00 kg (1.00 L of water has a mass of 1.00 kg)
Refer to Table 1.2.1 on page 12 for the relevant specific-heat (c) values.	$c_{iron} = 440\ \text{J kg}^{-1}\ \text{K}^{-1}$ $c_{water} = 4180\ \text{J kg}^{-1}\ \text{K}^{-1}$
Total energy lost by iron = total energy gained by water That is, the energy change, ΔQ, is equal for the copper and the water.	$\Delta Q_{iron} = \Delta Q_{water}$ $m_{iron}\ c_{iron}\ \Delta T_{iron} = m_{water}\ c_{water}\ \Delta T_{water}$
Substitute the known values, expand and simplify to solve for the initial temperature of the iron.	$m_{iron}\ c_{iron}\ \Delta T_{iron} = m_{water}\ c_{water}\ \Delta T_{water}$ $0.0500 \times 440 \times (T_{iron} - 17.0) = 1.00 \times 4180 \times (17.0 - 15.0)$ $22.0T_{iron} - 374 = 8360$ $22.0T_{iron} = 8734$ $T_{iron} = \frac{8734}{22.0}$ $= 397°C$

Worked example: Try yourself 1.4.2

CALCULATING THERMAL EQUILIBRIUM

A 75.0 g piece of copper is heated over a flame for several minutes. The iron is then plunged into an insulated, closed container containing 0.500 L of cool water, originally at 20.0°C. When thermal equilibrium is reached, the temperature of the water is found to be 22.0°C. If no water changes state to become steam and there are no other energy losses, then what was the temperature of the iron just before it was immersed in the water?

Solving 'heating' questions

Problems involving thermal energy often require calculations of both latent and specific heat as materials or systems change temperature and state during heating and/or cooling. They can require a number of steps and the process can at first seem quite complex. A flow diagram (Figure 1.4.1) often helps in understanding the steps and the form of heating or cooling required for each step.

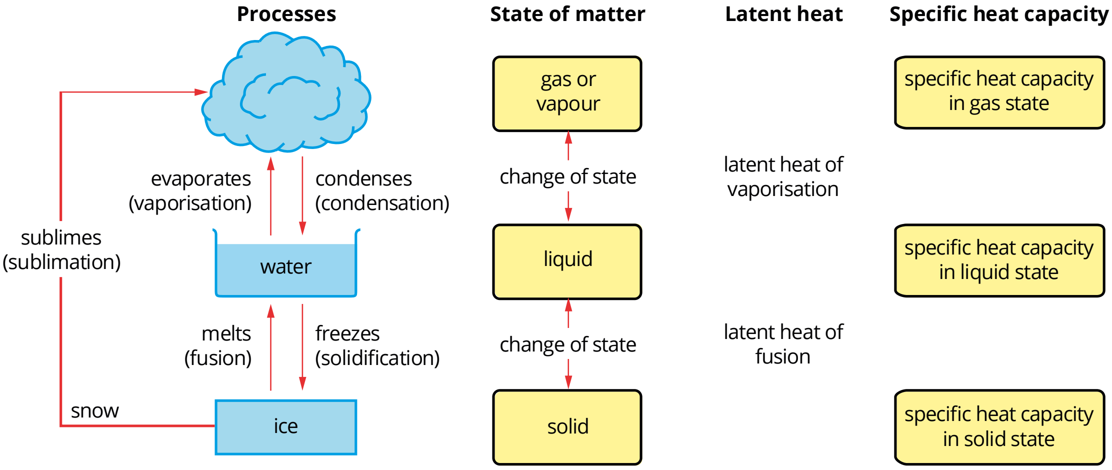


FIGURE 1.4.1 Solving heating and cooling questions involving changes of state.

Worked example 1.4.3

CHANGES OF STATE

Calculate the thermal energy required, in MJ, to convert 5.00 kg of ice at –20.0°C into steam at 100.0°C.

Thinking	Working
Identify the steps involved in the process.	Step 1: Ice at –20.0°C to ice at 0.00°C Step 2: Ice at 0.00°C to water at 0.00°C Step 3: Water at 0.00°C to water at 100.0°C Step 4: Water at 100.0°C to steam at 100.0°C
Identify values for L and c for each step. Use tables 1.2.1, 1.3.1 and 1.3.2 on pages 12, 16 and 17 to look up the values.	$c_{ice} = 2100\,J\,kg^{-1}\,K^{-1}$ $c_{water} = 4180\,J\,kg^{-1}\,K^{-1}$ $L_{fusion} = 3.34 \times 10^5\,J\,kg^{-1}$ $L_{vapour} = 22.5 \times 10^5\,J\,kg^{-1}$
Calculate the energy required for each step separately using the appropriate equation for specific heat or latent heat.	Step 1: Heating the ice $Q_1 = mc\Delta T = 5.00 \times 2100 \times 20.0 = 2.10 \times 10^5\,J$ Step 2: Melting the ice $Q_2 = mL_{fusion} = 5.00 \times 3.34 \times 10^5 = 1.67 \times 10^6\,J$ Step 3: Heating the water $Q_3 = mc\Delta T = 5.00 \times 4180 \times 100.0 = 2.09 \times 10^6\,J$ Step 4: Vapourising the water $Q_4 = mL_{vapour} = 5.00 \times 22.5 \times 10^5 = 1.125 \times 10^7\,J$
Add the energy required for each step together to find the total energy required.	$Q_T = Q_1 + Q_2 + Q_3 + Q_4$ $= (2.10 \times 10^5) + (1.67 \times 10^6) + (2.09 \times 10^6) + (1.125 \times 10^7)$ $= 1.52 \times 10^7\,J = 15.2\,MJ$

Worked Example: Try yourself 1.4.3

CHANGES OF STATE

Calculate the heat energy that must be lost, in J, to convert 5.00 kg of water vapour at 140.0°C into solid ice at 0.00°C.

1.4 Review

SUMMARY

- During a transfer of thermal energy, the materials in contact will eventually come to thermal equilibrium.
- The equilibrium temperature will depend upon the amount of energy transferred, the mass of the individual materials involved and the specific heat capacity of each material.
- Heating questions involving changes of state can be solved by breaking the question into parts or steps where there is either heating or cooling (specific heat change), or a change of state (latent heat change).

KEY QUESTIONS

1 Which absorbs the most thermal energy: 10 kg of iron changing from 10.0°C to 20.0°C, or 10 kg of aluminium changing from 10.0°C to 20.0°C? The specific heat of iron is 440 J $kg^{-1} K^{-1}$ and of aluminium is 900 J $kg^{-1} K^{-1}$.

2 10.0 kg of water at a temperature of 65.0°C is added to a bath containing 80.0 kg of water initially at 15.0°C. Calculate the final equilibrium temperature in degrees Celsius. The specific heat of water is 4180 J $kg^{-1} K^{-1}$.

3 A 20.0 kg block of copper at 100.0°C is put into a large pot containing 5.00 kg of water at 20.0°C. Assuming no energy is lost to the surrounding environment and no water changes state, calculate the final temperature of the mixture. Use the specific heat of water as 4180 J $kg^{-1} K^{-1}$ and copper as 390 J $kg^{-1} K^{-1}$.

4 Experts recommend that the bath temperature for a newborn baby should be 36.0°C. A group of students want to check what mass of water at 45.0°C they should add to 12.0 kg of water at 19.0°C in order to reach a final temperature suitable for the baby. Assuming no heat is lost to the bath or the surroundings, calculate the mass of warm water required.

5 A large 598 kg iron rod is produced in an iron smelter at 1250°C. It is cooled before transport by placing it in a water tank containing 938 litres of water at 21.0°C. Calculate the final temperature of the water and the iron in the water tank, assuming no heat is lost to the tank or the surroundings and the specific heat capacity of iron is 440 J $kg^{-1} K^{-1}$. Assume no water changes state.

6 A 10 kg block of iron at 20.0°C and a 10 kg block of aluminium at 20.0°C were dropped into a trough of 100 litres of water at 12.0°C. Calculate the final equilibrium temperature assuming no losses to the surrounding environment and no water changes state. Use c_{iron} = 440 J $kg^{-1} K^{-1}$, $c_{aluminium}$ = 900 J $kg^{-1} K^{-1}$ and c_{water} = 4180 J $kg^{-1} K^{-1}$.

7 50.0 kg of water at 20.0°C is placed in a large electric urn and heated to boiling point. Later, when the urn is checked, it is found that the water had boiled at this temperature until it had completely vaporised. Calculate how much energy was required to heat and boil away all of the water. The specific heat of water is 4180 J $kg^{-1} K^{-1}$ and the latent heat of vaporisation of water is 2.25×10^6 J kg^{-1}.

8 A chef uses a steamer to cook 3.00 kg of potatoes for a dish. The steam condenses on the potatoes and then the hot water continues to heat them up. Calculate the mass of steam at 100°C required to raise the temperature of the potatoes from 12.5°C to 85.0°C. The specific heat of water is 4180 J $kg^{-1} K^{-1}$ and the specific heat of potatoes is 3430 J $kg^{-1} K^{-1}$. The latent heat of vaporisation of water is 2.25×10^6 J kg^{-1}. Assume all of the steam condenses and no heat is lost to the surrounds.

9 A jeweller wants to melt a 1.25 kg ingot of silver metal by heating it from 20.0°C to its melting point of 961°C, before pouring it into moulds. Calculate the energy required to carry out the procedure if the specific heat of solid silver is 233 J $kg^{-1} K^{-1}$ and the latent heat of fusion of silver is 1.11×10^3 J kg^{-1}.

10 Steam cleaners use the extra energy contained in steam to break down grease and dirt from surfaces. A steam cleaner condenses 755 g of steam at 110°C onto a dirty surface which then ends up as water at 25.0°C. Calculate the energy transferred to the surface from the steam assuming no energy is transferred to any other substance. The specific heat capacity of steam is 2000 J $kg^{-1} K^{-1}$, the specific heat capacity of water is 4180 J $kg^{-1} K^{-1}$ and the latent heat of vaporisation of water to steam is 2.25×10^6 J kg^{-1}.

Chapter review

01

KEY TERMS

absolute zero
absorption
conduction
conductor
convection
electromagnetic spectrum
emission
emit
evaporation
frequency
heat
incident
insulator
internal energy
internal kinetic energy
internal potential energy
kelvin
kinetic particle model
latent heat
latent heat of fusion
latent heat of vaporisation
radiation
specific heat capacity
temperature
thermal contact
thermal energy
thermal equilibrium
volatile
wavelength
work

1 According to the kinetic particle model, which of the following is true for particles of matter?
  A They are in constant motion.
  B They have different sizes.
  C They have different shapes.
  D They are always floating.

2 What does the kelvin scale measure?

3 How does temperature differ from heat?

4 Convert:
  a 5°C to kelvin
  b 200 K to °C.

5 The kelvin scale is an example of an absolute scale. The Celsius scale is an arbitrary scale. List two key attributes of an absolute scale.

6 The Celsius scale has two fixed points—the freezing and boiling points of water—divided into 100 intervals. Explain why the Celsius scale cannot be considered an absolute scale.

7 The specific heat capacity of iron is approximately half that of aluminium. A ball of iron and a ball of aluminium, both at 80°C, are dropped into a thermally insulated jar that contains a mass of water, equal to that of the balls, at 20°C. Thermal equilibrium is eventually reached. Describe the final temperatures of each of the metal balls.

8 Two cubes, one silver and one iron, have the same mass and temperature. A quantity Q of heat is removed from each cube. Which one of the following properties causes the final temperatures of the cubes to be different?
  A density
  B specific heat capacity
  C latent heat of vaporisation
  D volume

9 A solid substance is heated but its temperature does not change. Explain what is occurring.

10 Which possesses the greater internal energy—1 kg of water boiling at 100°C or 1 kg of steam at 100°C? Explain why.

11 A liquid is evaporating, causing the liquid to cool. Explain why the temperature of the liquid decreases.

12 A 2.00 kg metal object requires 5.02×10^3 J of heat to raise its temperature from 20.0°C to 40.0°C. What is the specific heat capacity of the metal in $\text{J kg}^{-1}\text{K}^{-1}$? Give your answer to the nearest whole number.

13 How many joules of energy are required to melt exactly 80 g of silver? ($L_{\text{fusion}} = 0.88 \times 10^5\ \text{J kg}^{-1}$) Give your answer to two significant figures.

14 Copper is considerably more expensive than iron but is the preferred material for hot water pipes around the home while galvanised iron is generally used for cold water. Explain why copper is the preferred material for hot water pipes.

15 An insulated container holding 4.55 kg of ice at 0.00°C has 2.65 MJ of mechanical work done on it, while a heater provides 14 600 J of heat to the ice. If the latent heat of fusion of ice is $3.34 \times 10^5\ \text{J kg}^{-1}$, calculate the final temperature of the water. Assume that the increase in internal energy is first due to an increase in the potential energy and then an increase in the kinetic energy.

16 100 mL of water at 60°C is placed into a copper cup, also at 60°C and weighing 200 g. How much ice would be required to cool the water to 20°C?

17 A barista uses steam to heat up 425 g of milk with a specific heat capacity of $3930\ \text{J kg}^{-1}\text{K}^{-1}$ from 4.00°C to 70.0°C. Calculate the mass of steam, in kilograms, required to heat the milk if the steam starts at 100°C and ends up at the equilibrium temperature of the milk.

18 A student has a part-time job in a shop making blended icy fruit drinks. A 468 g lemon juice drink at 20.0°C, with a specific heat capacity of $3850\,J\,kg^{-1}\,K^{-1}$, has ice at 0.00°C added to it. Calculate the mass of ice required for the mixture to end up at 3.00°C. Use latent heat of fusion of ice as $3.34 \times 10^5\,J\,kg^{-1}$.

19 An airport worker uses a steam gun to melt the ice off the wheels of a plane in wintry Canada prior to departure. Calculate the mass of steam at 115°C required to convert 2.50 kg of ice at –12.5°C into water at 55°C. The specific heat capacity of steam is $2000\,J\,kg^{-1}\,K^{-1}$, of water is $4180\,J\,kg^{-1}\,K^{-1}$ and of ice is $2100\,J\,kg^{-1}\,K^{-1}$. The latent heat of fusion of water is $3.34 \times 10^5\,J\,kg^{-1}$ and the latent heat of vaporisation of water is $2.25 \times 10^6\,J\,kg^{-1}$.

20 1.50 kg of water at 22.0°C is poured on to an 18.0 kg iron barbeque hotplate at 545°C to cool it down. If all of the water is converted to steam at 100°C, calculate the final temperature of the hotplate. The specific heat capacity of water is $4180\,J\,kg^{-1}\,K^{-1}$ and of iron is $440\,J\,kg^{-1}\,K^{-1}$. The latent heat of vaporisation of water is $2.25 \times 10^6\,J\,kg^{-1}$.

CHAPTER 02

Moving heat around

In this chapter you will learn more about thermodynamic principles. You will learn about work as a change in energy forms that results in moving an object over a distance. You will also learn how thermal energy has the capacity to do mechanical work and how the internal energy of a system changes when it receives or transfers energy through heating or work.

This chapter also analyses the three processes through which thermal energy transfer happens within and across objects:

- conduction
- convection
- radiation.

Science as a Human Endeavour

- passive solar design for heating and cooling of buildings
- the development of the refrigerator over time
- the use of the sun for heating water
- engine cooling systems in cars.

Science Understanding

- a system with thermal energy has the capacity to do mechanical work [to apply a force over a distance]; when work is done, the internal energy of the system changes
- because energy is conserved, the change in internal energy of a system is equal to the energy added by heating, or removed by cooling, plus the work done on or by the system
- heat transfer occurs between and within systems by conduction, convection and/or radiation
- energy transfers and transformations in mechanical systems always result in some heat loss to the environment, so that the usable energy is reduced and the system cannot be 100 percent efficient. This includes applying the relationship:

$$\text{efficiency } (\eta) = \frac{\text{energy output}}{\text{energy input}} \times \frac{100\%}{1}$$

2.1 Work and efficiency

Most traditional power plants, such as that illustrated in Figure 2.1.1, use thermal energy to produce steam. The thermal energy comes from the burning of coal or the decay of a radioactive isotope. The steam produced does work turning turbines, which generate electricity. The capacity of thermal energy to do mechanical work is an important aspect in the study of heating processes.

FIGURE 2.1.1 Most power plants use thermal energy to produce steam. That steam does work turning turbines, which generate electricity.

ENERGY AND WORK

Energy is difficult to define, but can be described as the ability to do work. That means that energy has the potential to move objects over a distance.

Energy takes many different forms—for example, thermal, kinetic or potential. Despite the apparent different nature of the various forms of energy, any form of energy can be changed from one form to another. In order for any energy transformation to occur, say from heat to motion, work must be done. The work done on an object can be measured and, therefore, different types of energy can be compared by the work that they can do.

The work (W) done by a system, measured in joules (J), is calculated by multiplying the force of magnitude F(N) applied to an object that moves it a distance s(m) in the direction of the force:

$$W = Fs$$

Mechanical energy to thermal energy

The experiments of James Prescott Joule (1818–89) were fundamental in understanding how mechanical work could transfer mechanical energy into thermal energy. Joule noticed that stirring water could cause its temperature to rise. He designed a way to measure the relationship between the energy transferred when stirring the water and the change in temperature. A metal paddle wheel was rotated by falling masses and this churned the water around in an insulated can. Joule's original experimental set-up is shown in Figure 2.1.2.

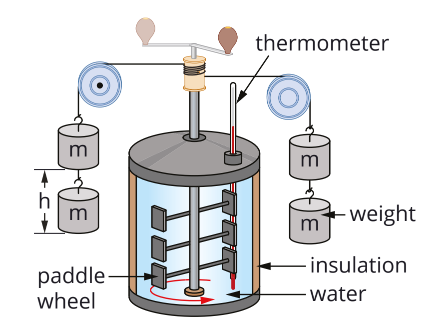


FIGURE 2.1.2 Joule's apparatus for investigating the heating effect of mechanical work allowed the transfer of heat energy to be measured and related to other forms of energy.

The work done in the system used in Joule's experiment was calculated by multiplying the weight of the falling masses (which churned the water) by the distance they fell. The heat generated was calculated from the mass of the water and the temperature rise. Joule found that the same amount of mechanical work ($W = Fs$) always produced exactly the same amount of heating. This meant that heat was another form of energy, known now as thermal energy. Joule found that approximately 4.18 J of work was needed to raise the temperature of 1 g of water by 1°C, a figure now referred to as the specific heat capacity of water ($c = 4180\,\text{J kg K}^{-1}$). The heat energy transferred (Q) by the transfer of mechanical energy to a mass m for a change in temperature of ΔT is thus equal to $Q = mc\Delta T$.

Thermal energy to mechanical energy

A system with thermal energy has the capacity to do mechanical work; that is, to apply a force that moves something. Through work, energy is transferred from one system to another. The system doing work will lose internal energy; the system on which work is done will gain **internal energy**. The change in internal energy of a system is equal to the energy added by heating or removed by cooling, plus the work done on or by the system.

The most common examples of thermal energy being used to do work are thermal power stations. Almost all traditional power plants (Figure 2.1.3) are thermal, including nuclear, coal, solar thermal electric and many natural gas power plants.

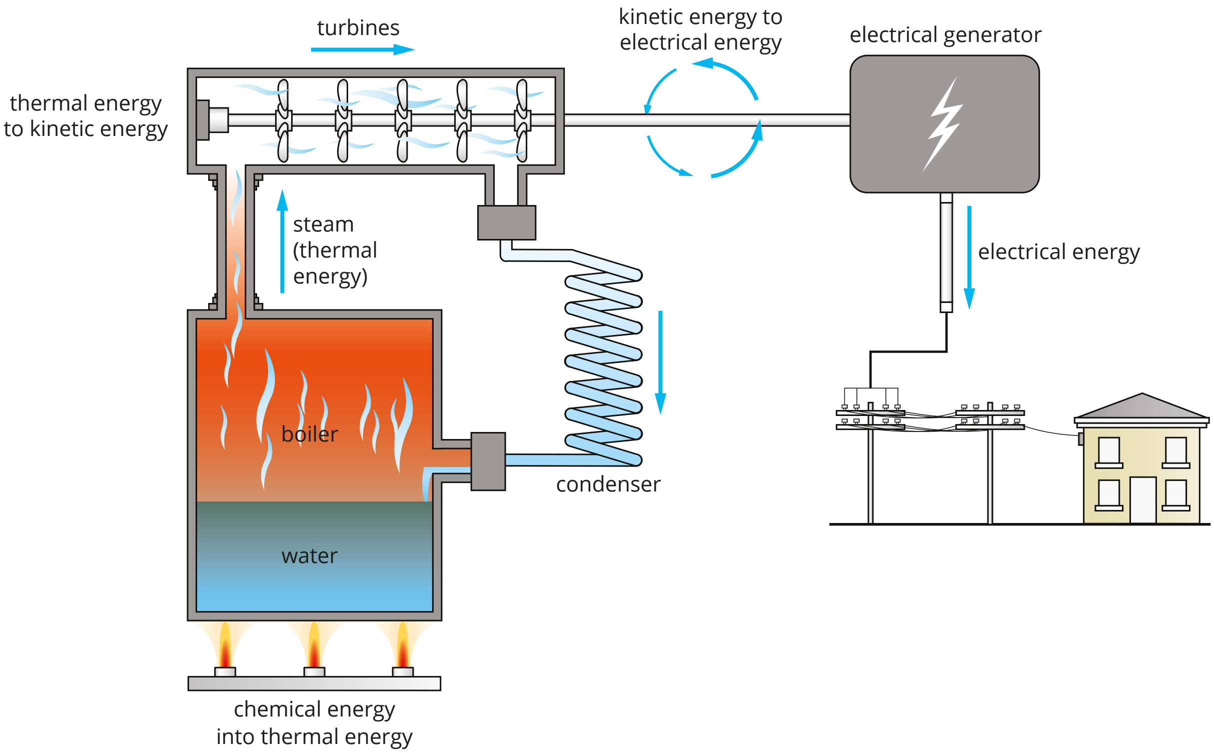


FIGURE 2.1.3 A thermal power station, which uses steam to drive turbines and generate electricity.

A thermal power station is a power plant that uses fuel to heat water to steam under high pressure. The steam performs mechanical work on steam turbines, which drive electrical generators. After steam passes through the turbine, it is condensed in a condenser and recycled to where it was heated.

Through this process, thermal power stations convert forms of heat energy into electrical energy. The greatest variation in the design of thermal power stations is due to the different fuels (mostly fossil fuels) used to heat the water.

Efficiency of energy transformations

In the real world, energy transformations, such as the ones described for a thermal power plant, are never perfect—there is always some energy lost. Because of this, for a system to continue operating (doing work), it must be constantly provided with energy. The percentage of energy that is effectively transformed by a system is called the efficiency of that device. A device operating at 45% efficiency converts 45% of its supplied energy into the useful new form. The other 55% is lost to the surroundings, usually as heat and/or sound. Table 2.1.1 shows the efficiencies of some devices.

The efficiency of a transfer from one energy form to another is expressed as:

$$\text{efficiency } (\eta) = \frac{\text{useful energy transferred}}{\text{total energy supplied}} \times 100\%$$

$$= \frac{\text{energy output}}{\text{energy input}} \times 100\%$$

TABLE 2.1.1 Approximate efficiencies of some common things.

Device	Energy transfer	Efficiency (%)
electric motor	electric to kinetic	90
gas heater	chemical to thermal	75
incandescent light globe	electric to light	2
compact fluorescent light	electric to light	10
LED household light	electric to light	15
steam turbine	thermal to kinetic	45
coal-fired generator	chemical to electrical	30
high-efficiency solar cell	radiation to electrical	35
car engine	chemical to kinetic	25
open fireplace	chemical to thermal	15
human body	chemical to kinetic	25

Worked example 2.1.1

ENERGY EFFICIENCY

The energy input of a gas-fired power station is 1100 MJ. The electrical energy output is 300 MJ. What is the efficiency of the power station?

Thinking	Working
Recall the formula for efficiency of energy transfers.	$\text{efficiency } (\eta) = \frac{\text{energy output}}{\text{energy input}} \times 100\%$
Substitute the known values into the formula. Both values are in MJ, so there's no need to change them to J.	output = 300 MJ input = 1100 MJ $\text{efficiency } (\eta) = \frac{300}{1100} \times 100$
Solve the equation.	efficiency = 27%

Worked example: Try yourself 2.1.1

ENERGY EFFICIENCY

An electric kettle uses 23.3 kJ of electrical energy as it boils a quantity of water. The efficiency of the kettle is 18%. How much electrical energy is used in actually boiling the water? Give your answer in kJ.

2.1 Review

SUMMARY

- When work is done, the internal energy of a system changes.
- A system with thermal energy has the capacity to do mechanical work.
- The mechanical work, W, done by a system is calculated by multiplying the force of magnitude F applied to an object that moves it a distance s in the direction of the force:

  $W = Fs$
- The efficiency of a transfer from one energy form to another is given by:

$$\text{efficiency } (\eta) = \frac{\text{useful energy transferred}}{\text{total energy supplied}} \times 100\%$$

$$= \frac{\text{energy output}}{\text{energy input}} \times 100\%$$

KEY QUESTIONS

1 Identify all of the energy transformations that take place in a coal-fired electrical generator.

2 What caused a rise in temperature in the water in Joule's experiment?

3 How much work is done on an object of 4.5 kg when it is lifted vertically at a constant speed through a displacement of 6.0 m? (Use $g = 9.80\,\text{m}\,\text{s}^{-2}$.)

4 A student places a heating element and a paddle wheel apparatus in an insulated container of water. She calculates that the heater adds 2530 J of heat energy to the water, and the paddle wheel does 240 J of work on the water. Calculate the change in internal energy of the water.

The following information applies to questions 5 and 6

A weightlifter raises a 100 kg mass 2.4 m above the ground in a weightlifting competition. After holding it for 3.0 s he places it back on the ground.

5 How much work has been done by the weightlifter in raising the mass? Use $g = 9.80\,\text{m}\,\text{s}^{-2}$ and give your answer to two significant figures.

6 How much additional work is done during the 3.0 s he holds it steady?

7 A particular model of reverse-cycle air conditioner produces 1.2 kW of useful heat from 4.8 kW of electrical energy. What is the efficiency of the air conditioner?

8 An electric drill uses 3.6 kJ of energy to drill a hole through a sheet of steel. The efficiency of the drill itself is 70%. How much electrical energy, in kJ, is used in actually making the hole?

9 A cook uses an egg beater on a warm chocolate sauce. If the cook does 845 J of work on the sauce while the warm sauce loses 1239 J of heat to the environment, what is the change in internal energy of the chocolate sauce?

10 A coal-fired generator has an efficiency of approximately 30%. If 2000 J of energy is supplied to the generator, then how much is converted into electrical energy?

2.2 Conduction

FIGURE 2.2.1 Emperor penguin chicks avoid heat loss through conduction by sitting on the adult's feet. In this way they avoid contact with the ice.

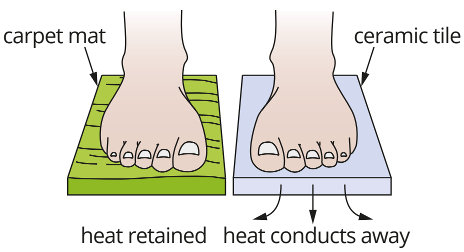


FIGURE 2.2.2 Ceramic floor tiles are good conductors of heat. They conduct heat away from the foot readily and so your feet feel cold on tiles. The carpet mat is a thermal insulator. Thermal energy from the foot is not transferred away as quickly and so your foot doesn't feel as cold.

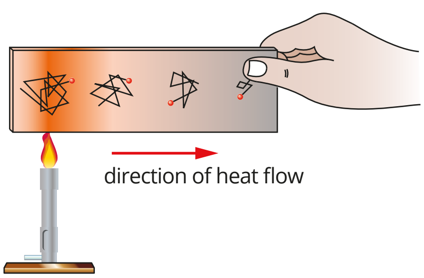


FIGURE 2.2.3 Thermal energy is passed on by collisions between adjacent particles.

If two objects are at different temperatures and are in thermal contact (that is, they can exchange energy via heat processes), then thermal energy will transfer from the hotter object to the cooler object. Figure 2.2.1 shows how, by preventing the chick's thermal contact with the cold ice, this adult penguin is able to protect the vulnerable penguin offspring.

This section focuses on heat transfer by conduction.

CONDUCTORS AND INSULATORS

Conduction is the process by which heat is transferred from one place to another without the net movement of particles (atoms or molecules). Conduction can occur within a material or between materials that are in thermal contact. For example, if one end of a steel rod is placed in a fire, heat will travel along the rod so that the far end of the rod will also heat up; or if a person holds an ice cube, then heat will travel from their hand to the ice.

While all materials will conduct heat to some extent, this process is most significant in solids. It is important in liquids but plays a lesser role in the movement of energy in gases.

Materials that conduct heat readily are referred to as good **conductors**. Materials that are poor conductors of heat are referred to as insulators. An example of a good conductor and a good **insulator** can be seen in Figure 2.2.2.

In secondary physics, the terms 'conductor' and 'insulator' are used in the context of both electricity and heating processes. What makes a material a good conductor of heat doesn't necessarily make it a good conductor of electricity. The two types of conduction are related but it's important not to confuse the two processes. A material's ability to conduct heat depends on how conduction occurs within the material.

Conduction can happen in two ways:

- energy transfer through molecular or atomic collisions
- energy transfer by free electrons.

THERMAL TRANSFER BY COLLISION

The kinetic particle model explains that particles in a solid substance are constantly vibrating within the material structure and so interact with neighbouring particles. If one part of the material is heated, then the particles in that region will vibrate more rapidly. Interactions with neighbouring particles will pass on this kinetic energy throughout the system via the bonds between the particles (Figure 2.2.3).

The process can be quite slow since the mass of the particles is relatively large and the vibrational velocities are fairly low. Materials for which this method of conduction is the only means of heat transfer are likely to be poor conductors of heat or even thermal insulators. Materials such as glass, wood and paper are poor conductors of heat.

THERMAL TRANSFER BY FREE ELECTRONS

Some materials, particularly metals, have electrons that are not directly involved in any one particular chemical bond. Therefore, these electrons are free to move throughout the lattice of positive ions.

If a metal is heated, then not only will the positive ions within the metal gain extra energy but so will these free electrons. As the electron's mass is considerably less than the positive ions, even a small energy gain will result in a very large gain in velocity. Consequently, these free electrons provide a means by which heat can be quickly transferred throughout the whole of the material. It is therefore no surprise that metals, which are good electrical conductors because of these free electrons, are also good thermal conductors.

THERMAL CONDUCTIVITY

Thermal conductivity describes the ability of a material to conduct heat. It is temperature dependent and is measured in watts per metre per kelvin ($W m^{-1} K^{-1}$). Table 2.2.1 highlights the difference in conductivity in metals compared with other substances.

TABLE 2.2.1 Thermal conductivities of some common materials.

Material	Conductivity ($W m^{-1} K^{-1}$)
silver	420
copper	380
aluminium	240
steel	60
ice	2.2
brick, glass	≈ 1
concrete	≈ 1 (depending on composition)
water	0.6
human tissue	0.2
wood	0.15
polystyrene	0.08
paper	0.06
fibreglass	0.04
air	0.025

Factors affecting thermal conduction

The rate at which heat is transferred through a system depends on the:

- nature of the material. The larger a material's thermal conductivity, the more rapidly it will conduct heat energy.
- temperature difference between the two objects. A greater temperature difference will result in a faster rate of energy transfer.
- thickness of the material. Thicker materials require a greater number of collisions between particles or movement of electrons to transfer energy from one side to the other.
- surface area. Increasing the surface area relative to the volume of a system increases the number of particles involved in the transfer process, increasing the rate of conduction.

The rate at which heat is transferred is measured in joules per second ($J s^{-1}$), or watts (W).

PHYSICSFILE

Igloos

It seems strange that an igloo can keep a person warm when ice is so cold. Igloos are constructed from compressed snow that contains many air pockets. The air in these pockets is a poor conductor of heat, which means heat inside the igloo is not easily transferred away. The body heat of the occupant, as well as that of his or her small heat source, is trapped inside the igloo and is able to keep them warm.

FIGURE 2.2.4 Air pockets in compressed snow enable igloos to keep their occupants warm relative to outside temperatures.

EXTENSION

Thermal conductivity

The rate of energy transfer by conduction (energy per unit time) through a material can be calculated using:

$$\frac{Q}{t} = \frac{kA\Delta T}{L}$$

where $\frac{Q}{t}$ is the heat energy, Q, transferred in joules (J) per unit time, t, in seconds (s)

k is the thermal conductivity of the material ($W m^{-1} K^{-1}$)

A is the surface area perpendicular to the direction of heat flow, in metres squared (m^2)

ΔT is the temperature difference across the material in kelvin or degrees Celsius (K or °C)

L is the thickness of the material through which the heat is being transferred, in metres (m).

Designers use this relationship when calculating the insulating ability of the fill inside clothing such as parkas. Architects and builders use it to calculate the efficiency of building insulation.

Guidelines exist to ensure the efficiency of insulating materials. Building materials that limit the transfer of heat help to keep houses warm in winter and cool in summer. This saves money and helps to reduce carbon dioxide emissions from the use of gas or electricity to heat houses.

PHYSICS IN ACTION

Passive solar design

Rising energy costs and dwindling resources have caused many countries to try to reduce the amount of fossil fuels used to provide energy for our homes, schools and industry.

Passive heating of a building through passive solar design is one way to reduce the overall energy being consumed within the building. Buildings that are designed with passive solar heating and cooling in mind take advantage of natural climate to maintain a comfortable temperature indoors. Figure 2.2.5 describes some of the principles of a passive solar house.

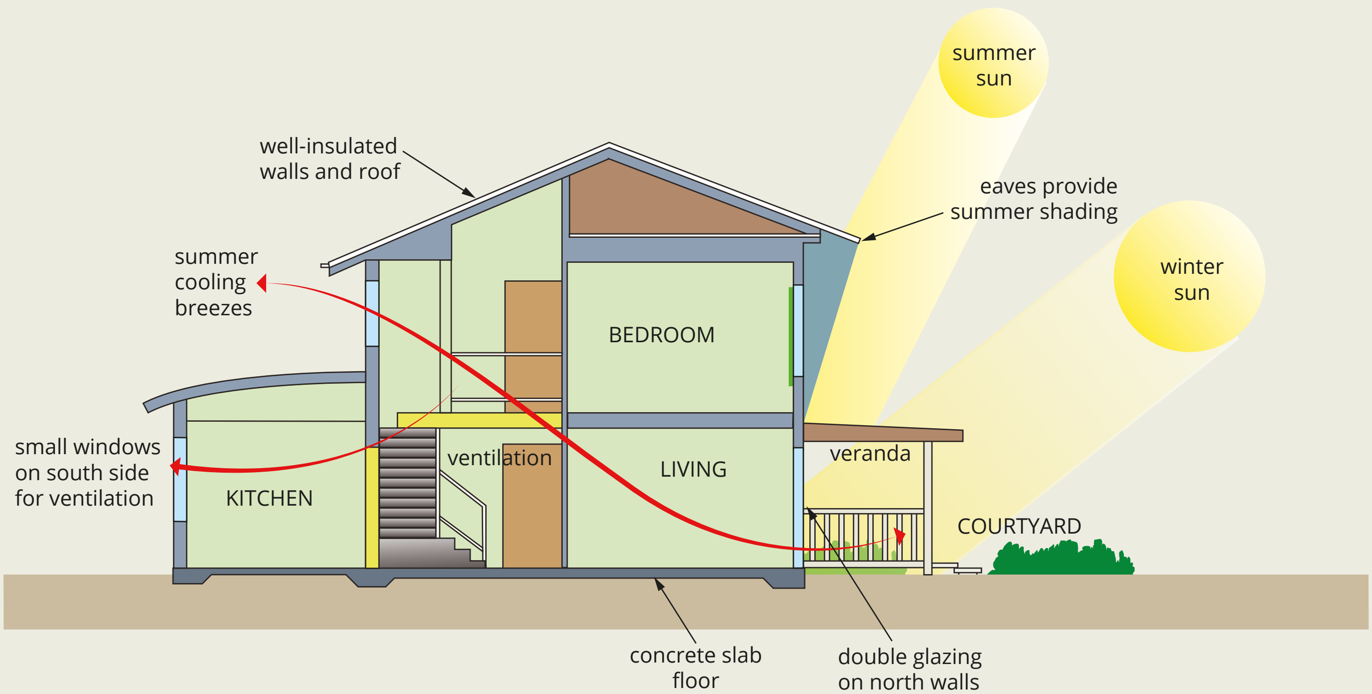


FIGURE 2.2.5 Aspects of good passive solar design applied to housing design can lead to significant energy savings.

Insulation prevents loss by conduction in winter and gain of heat by conduction in summer. Solid brick walls and concrete slab floors provide large amounts of thermal mass, which take longer to heat up on hot days and cool down on cold nights.

Positioning large windows on the northern side of a house allows generous amounts of solar energy to warm the interior in winter. Overhanging eaves and deciduous trees provide shade in summer. Small windows to the south provide ventilation but reduce heat loss on the shady side of the house. Blank walls heavily shaded by trees or overhangs avoid the sun of hot summer afternoons and shield against the prevailing cold westerlies of winter.

Siting living areas on the northern side allows the winter sun to warm areas of the house where it's needed during the day. Rooms that are used little during the day can be sited on the southern side of the house.

Heavy curtains and double-glazing (Figure 2.2.6) can reduce conductive heat loss at night and in winter. They also reduce solar energy gains in summer. Conventional single glazing allows heat to be lost through conduction. Double glazing provides additional thermal resistance by adding a sealed space between two glass panes. The air gap conducts much less heat. Increasingly, argon gas is used to fill the space between the panes instead of air, because it has a lower conductivity than air. Good use of double glazing can reduce heat loss or gain by at least 50%.

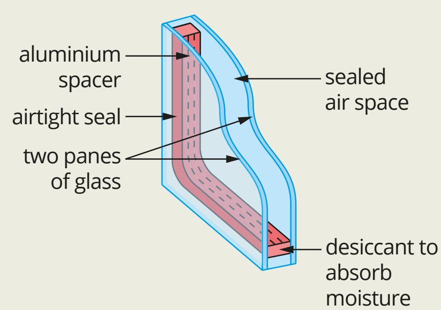


FIGURE 2.2.6 Double-glazed window construction reduces heat loss through windows that would otherwise transfer considerable energy through conduction.

2.2 Review

SUMMARY

- Conduction is the process of heat transfer within a material or between materials without the overall transfer of the substance itself.
- All materials will conduct heat to a greater or lesser degree. Materials that readily conduct heat are called good thermal conductors. Materials that conduct heat poorly are called thermal insulators.
- Whether a material is a good conductor depends on the method of conduction:
  - Heat transfer by molecular collisions alone occurs in poor to very poor conductors.
  - Heat transfer by molecular collisions and free electrons occurs in good to very good conductors.
- The rate of conduction depends on the temperature difference between two materials, the thickness of the material, the surface area and the nature of the material.

KEY QUESTIONS

1 Explain why the process of conduction by molecular collision is slow.

2 Why are metals more likely to conduct heat than wood?

3 List the properties of a material that affect its ability to conduct heat.

4 Stainless steel saucepans are often manufactured with a copper base. What is the most likely reason for this?

5 One way of making a house energy-efficient is to use double-glazed windows. These consist of two panes of glass with air trapped in between the panes. On a hot day, the energy from the hot air outside the house is not able to penetrate the air gap and so the house stays cool.
Which of the following best explains why double-glazing works?
A Air is a conductor of heat and so the thermal energy is able to pass through.
B Air is a conductor of heat and so the thermal energy is not able to pass through.
C Air is an insulator of heat and so the thermal energy is not able to pass through.
D Air is an insulator of heat and so the thermal energy is able to pass through.

6 How does a down-filled quilt keep a person warm in winter?

7 Fibreglass insulation batts are thick and lightweight, and they make a house more energy-efficient. On a cold July night, the external temperature in the roof of an insulated house is 6°C. The air temperature near the ceiling inside the house is 20°C. Explain how ceiling insulation decreases energy loss.

8 On a cold day, the plastic or rubber handles of a bicycle feel much warmer than the metal surfaces. Explain this in terms of the thermal conductivity of each material.

9 Explain how windows should be placed on homes to maximise solar gain.

10 What is the main difference between single- and double-glazed windows that makes double-glazed windows better thermal insulators?

2.3 Convection

Convection is the transfer of thermal energy within a fluid (liquid or a gas) by the movement of hot areas from one place to another. Unlike other forms of heat transfer such as conduction and radiation, convection involves the mass movement of particles within a system over a distance that can be quite considerable.

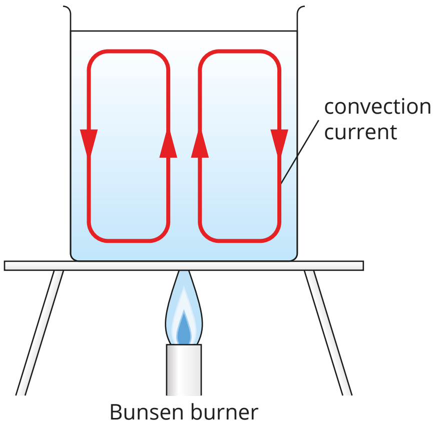


FIGURE 2.3.1 When a liquid or gas is heated, it becomes hotter and less dense so will rise. The colder, denser fluid will fall. As this fluid heats up, it in turn will rise, creating a convection current.

HEATING BY CONVECTION

Although liquids and gases are generally not very good conductors of thermal energy, heat can be transferred quite quickly through liquids and gases by convection. Unlike other forms of thermal energy transfer, convection involves the mass movement of particles within a system over a distance.

As a fluid is heated, the particles within it gain kinetic energy and push apart due to the increased vibration of the particles. This causes the density of the heated fluid to decrease and the heated fluid rises. Colder fluid, with slower moving particles, is more dense and heavier and hence falls, moving in to take the place of the warmer fluid. A convection current forms when there is warm fluid rising and cool fluid falling. This action can be seen in Figure 2.3.1. Upwellings in oceans, wind and weather patterns are at least partially due to convection on a very large scale.

It is difficult to quantify the thermal energy transferred via convection but some estimates can be made. The rate at which convection will occur is affected by:

- the temperature difference between the heat source and the convective fluid
- the surface area exposed to the convective fluid.

In a container, the effectiveness of convection in transferring heat depends on the placement of the source of heat. For example, the heating element in a kettle is always found near the bottom of the kettle. From this position, convection currents form throughout the water to heat it more effectively (Figure 2.3.2(a)). If the heating element were placed near the top of the kettle, convection currents would form only near the top. This is because the hotter water is less dense than the cooler water below and will remain near the top. Convection currents will not form throughout the water (Figure 2.3.2(b)).

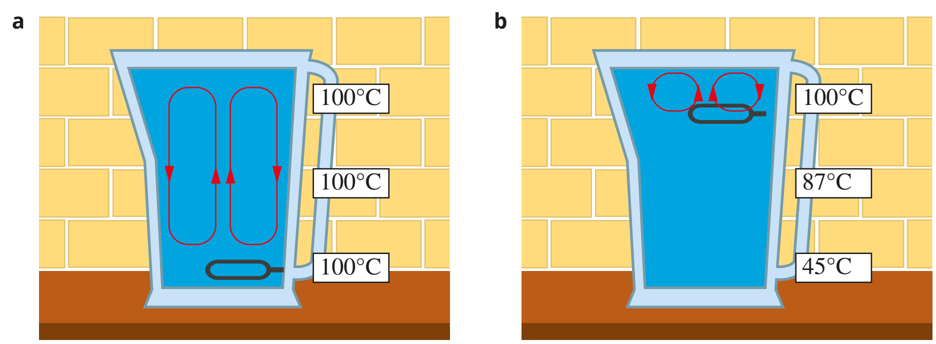


FIGURE 2.3.2 (a) By placing the heating element at the bottom of a kettle, the water near the bottom is heated and rises, forming convection currents throughout the entire depth of the water. (b) If the heating element is placed near the top of the kettle, the convection currents form near the top and heat transfer is slower.

There are two main causes of convection:

- forced convection, for example ducted heating in which air is heated and then blown into a room
- natural convection, such as that illustrated in Figure 2.3.2, when a fluid rises as it is heated.

A dramatic example of natural convection is the thunderhead clouds of summer storms (Figure 2.3.3), which form when hot, humid air from natural convection currents is carried rapidly upwards into the cooler upper atmosphere.

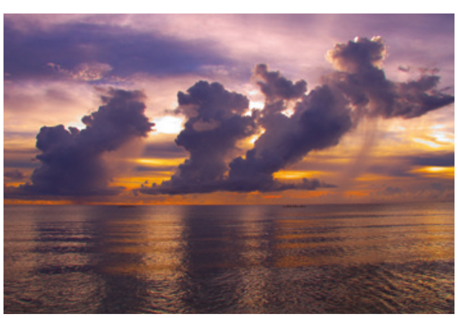


FIGURE 2.3.3 The thunderheads of summer storms are a very visible indication of natural convection in action.

PHYSICS IN ACTION

Wind chill

Convective effects are the main means of heat transfer that lead to the 'wind chill' factor. The wind blows away the thin layer of relatively still air near the skin that would normally act as a partial insulator in still air. Cooler air comes in closer contact with the skin and heat loss increases. It feels as if the 'effective' temperature of the surrounding air has decreased. Skiers can experience similar effects simply from the wind created by their own motion.

In cold climates the wind chill factor can become an important factor to consider. The chilling effect is even more dramatic when the body or clothing is wet, increasing evaporative cooling (Figure 2.3.4). Evaporation is the process where a liquid becomes a gas, and in the process a large amount of thermal energy is taken away from the remaining liquid. Bushwalkers look for clothing made from materials that dry rapidly after rain and which carry moisture from the perspiration of heavy exertion away from the skin.

FIGURE 2.3.4 In cold environments, such as Antarctica, wind chill factor greatly increases the chilling effect and hence the risk of hypothermia.

PHYSICSFILE

Natural convection in air

There is often a temperature difference between the land and the sea in coastal environments. The temperature of the water hardly changes between night and day due to its high specific heat capacity but the land can become much hotter through the day. As the air over the land is heated, it rises and is replaced by cooler, more dense air from over the sea (Figure 2.3.5). This moving air creates a sea breeze that is experienced in most coastal areas in Australia during summer, including the famous Fremantle Doctor. This makes living on the coast appealing during hot weather.

At night, the process is reversed. The land cools more quickly and so does the air above it. This denser air moves out over the water, displacing the now relatively lighter and warmer air over the sea, and a land breeze is created.

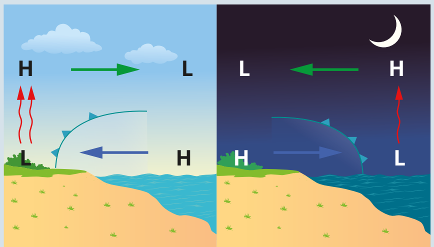


FIGURE 2.3.5 Land sea breezes are created by convection currents caused by the temperature difference between air over the sea and air over the land.

2.3 Review

SUMMARY

- Convection is the transfer of heat within a fluid (liquid or gas).
- Convection involves the mass movement of particles within a system over a distance.
- A convection current forms when there is warm fluid rising and cool fluid falling.

KEY QUESTIONS

1. Through what states of matter can convection occur?
2. In which direction does the transfer of heat in a convection current initially flow?
3. Pilots of glider aircraft or hang-gliders, some birds such as eagles and some insects rely on 'thermals' to give them extra lift. Explain how these rising columns of air are established.
4. Compare the conductive and convective abilities of liquids and gases.
5. Convection is a method of heat transfer through fluids. Explain whether it is possible for solids to pass on their heat energy by convection.
6. On a hot day, the top layer of water in a swimming pool can heat up while the lower, deeper parts of the water can remain quite cold. Explain, using the concept of convection, why this happens.
7. List the steps in the process by which thermal energy is transferred from one place to another within a liquid or gas.
8. What factors affect the rate of thermal convection?

2.4 Radiation

Both convection and conduction involve the transfer of heat through matter. Life on Earth depends upon the transfer of energy from the Sun through the near-vacuum of space. If heat could only be transferred by the action of particles, then the Sun's energy would never reach Earth. Radiation is a means of transfer of heat without the movement of matter.

RADIATION

In this context, **radiation** is a shortened term for electromagnetic radiation, which includes visible, ultraviolet and infrared light. Together with other forms of light, these make up the **electromagnetic spectrum**.

The transfer of heat from one place to another without the movement of particles is by electromagnetic radiation. Electromagnetic radiation travels at the speed of light. When electromagnetic radiation hits an object, it will be partially reflected, partially transmitted and partially **absorbed**. The absorbed part transfers thermal energy to the absorbing object and causes a rise in temperature. When you hold a marshmallow by an open fire, you are using radiation to toast the marshmallow, as shown in Figure 2.4.1.

FIGURE 2.4.1 Heat transfer from the flame to the marshmallow is an example of radiation.

Electromagnetic radiation is **emitted** by all objects that are at a temperature above absolute zero (0 K or −273°C). The **wavelength** and **frequency** of the emitted radiation depends on the internal energy of the object. As temperature is related to the internal energy of an object, the higher the temperature of the object, the higher the frequency and the shorter the wavelength of the radiation emitted. This can be seen in Figure 2.4.2.

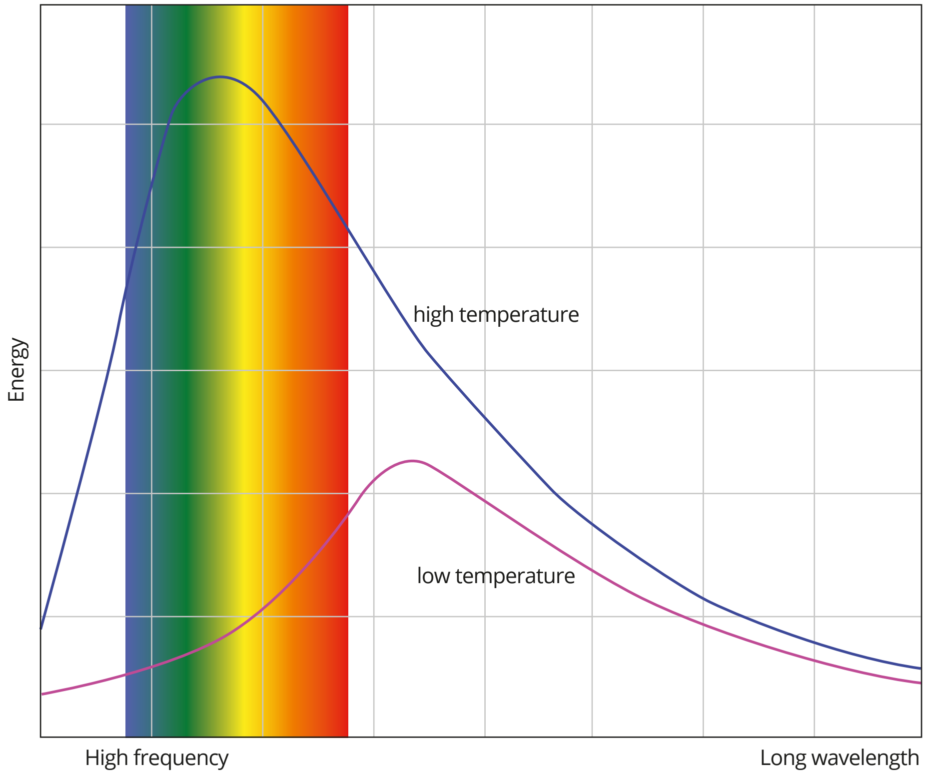


FIGURE 2.4.2 A system emits radiation over a range of frequencies. At a low temperature, it will emit small amounts of radiation of longer wavelengths. As the temperature of the system increases, more short-wavelength radiation is emitted and the total radiant energy emitted increases.

A human body emits radiation in the infrared range of wavelengths, while hotter objects emit radiation of a higher frequency and shorter wavelength. Hotter objects can emit radiation in the range of visible, ultraviolet and shorter wavelengths of the electromagnetic spectrum. For example, as a red-hot fire poker heats up further, it becomes yellow-hot.

EMISSION AND ABSORPTION OF RADIANT ENERGY

All objects both absorb and emit thermal energy by radiation. If an object absorbs more thermal energy than it emits, its temperature will increase. If an object emits more energy than it absorbs, its temperature will decrease. If no temperature change occurs, the object and its surroundings are in thermal equilibrium.

While all objects emit some radiation, they will not all emit or absorb at the same rate.

A number of factors affect both the rate of emission and the rate of absorption.

- Surface area—the larger the exposed surface area, the higher the rate of radiant transfer.
- Temperature—the greater the difference between the temperature of the absorbing or emitting surface and the temperature of the surrounding objects, the greater the rate of energy transfer by radiation.
- Wavelength of the **incident** radiation—matte black surfaces are almost perfect absorbers of radiant energy at all wavelengths. Highly reflective surfaces are good reflectors of all wavelengths. An example of how reflective surfaces can be exploited is shown in Figure 2.4.3. For all other surfaces, the absorption of particular wavelengths of radiant energy will be affected by the wavelength of that energy. For example, although white surfaces absorb visible wavelengths of radiant energy poorly, white surfaces will absorb infrared radiation just as well as black surfaces do.
- Surface colour and texture—the characteristics of the surface itself determine how readily that particular surface will emit or absorb radiant energy.

Matte black surfaces will absorb and emit radiant energy faster than shiny, white surfaces. This means that a roughened, dark surface will heat up faster than a shiny, light one. Matte black objects will also cool down faster since they will radiate energy just as efficiently as they absorb energy. Car radiators are painted black to increase the emission by radiation of the thermal energy that is collected from the car engine. A silvered surface will reflect radiant energy, taking longer to heat up.

FIGURE 2.4.3 The silvered surface of an emergency blanket reflects thermal energy back to the body, and retains the radiant energy, which would normally be lost. This simple method works as excellent thermal insulation.

PHYSICSFILE

Thermal imaging

All objects emit radiant energy. Humans are warm-blooded and emit radiation in the infrared region of the electromagnetic spectrum (Figure 2.4.4). This radiation is not visible to us, but can be detected using thermal imaging devices or night-vision goggles. These devices are used by the military for surveillance, and by search and rescue personnel. Some animals, notably some varieties of bugs, are able to detect infrared radiation.

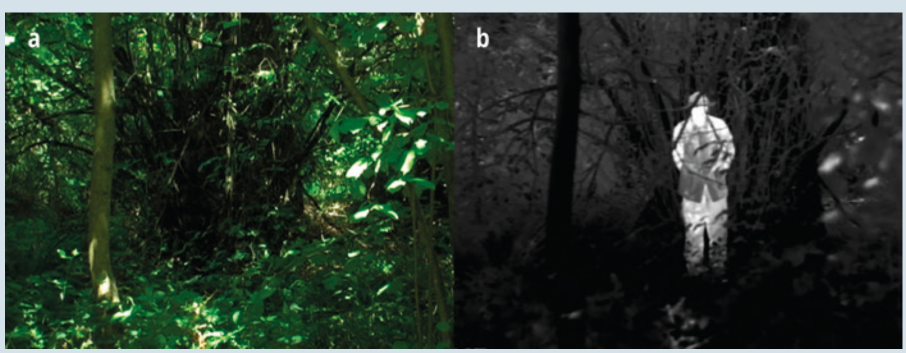


FIGURE 2.4.4 This person is difficult to see using the visible range of the electromagnetic spectrum (a), but they are warmer than the trees surrounding them and produce a stronger infrared emission, which the thermal imaging device is able to detect (b).

PHYSICSFILE

Heat loss in humans

It is estimated that, at normal room temperature, about 50 per cent of a person's heat loss is by radiation.

PHYSICS IN ACTION

Engine cooling systems in cars

Have you ever noticed the heat coming off a car engine after a journey? If not, next time you go for a drive, hold your hand over the engine at the end of the journey and notice how hot it is.

This is because a car engine is not 100% efficient. In fact, only about 20–30% of the chemical energy contained within the fuel is transformed into mechanical energy needed to drive the car. The rest of the energy released when the fuel burns inside a car engine is lost as thermal energy.

A car engine does need some heat to keep it running smoothly and reduce wear on the parts. The rest of the heat has to be removed to avoid the engine overheating. This thermal energy is removed by the car engine's cooling system.

A car's cooling system works by circulating a fluid, called a coolant, in pipes around the engine. This coolant absorbs some of the thermal energy generated during combustion of the fuel in the engine. As the hot coolant is pumped away from the engine, it removes thermal energy, regulating the engine temperature.

The hot coolant is then passed through a heat exchanger, called a radiator (Figure 2.4.5). The radiator is a system of narrow tubes, which give it a high surface area. A greater surface area means more energy can be radiated. Air from outside the car is blown across these pipes. This air can be moved across the radiator either by a fan or just by using the airflow over the moving car.

As the air is blown across the radiator, thermal energy is transferred from the coolant to the air. The air heats up and the coolant cools down. The coolant is then recirculated around the engine to continue the cooling process.

FIGURE 2.4.5 A radiator is often placed at the front of the car where it can take advantage of the airflow over the moving car.

PHYSICS IN ACTION

Solar water heating

The Sun's energy falls on the Earth's surface at the rate of about 3.6 MJ on each square metre per hour. This means that an average roof of around 200 m^2 receives 3600 MJ (or 3.6 GJ) of energy in 5 hours of sun. Given that an average household might use around 72 MJ of electrical energy in a day you can see the possibilities!

You can make use of this free source of energy by using solar cells (also called photovoltaic cells), which transform radiant energy from the Sun into electrical energy. Another much simpler way to harness the energy from the Sun is to use it to heat water.

If you have ever used a solar camp shower, you have taken advantage of solar hot water. A solar camp shower is black plastic container filled with water and hung out in the sunshine. Recall that darker surfaces will absorb more radiant energy than lighter surfaces. The black plastic absorbs radiant energy from the Sun and this energy is converted to thermal energy. This thermal energy heats the water inside.

The solar hot-water systems used in homes around Australia and countries such as Spain, China and India, work on similar principles. A solar collector panel is placed on the sunniest side of the roof and connected to a water storage tank.

The solar collector is usually made up of dark-coloured pipes that allow the flow of a fluid. In warmer climates, such as Australia, the fluid is regular drinking water, which becomes the hot water used within the home. In cooler climates, the fluid is often a mix of water and antifreeze, which stops the water freezing in the pipes. This water/antifreeze mixture is pumped through a water storage tank in closed pipes, where it transfers the thermal energy to the hot water for the house.

There are two main types of solar water heaters: passive and active. Active solar water heaters use electrical pumps to move the hot water around the system. Passive systems (Figure 2.4.6) rely on convection currents to move the hot water. The natural convective currents can be a low-energy means of moving hot water from a solar collector to a storage tank without the use of an electric pump. A simple passive system is shown the diagram below.

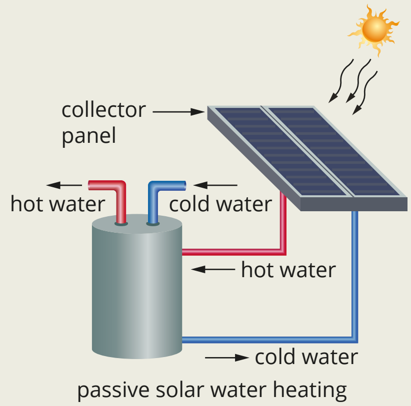


FIGURE 2.4.6 A simple passive solar hot-water system.

2.4 Review

SUMMARY

- Any object whose temperature is greater than absolute zero emits thermal energy by radiation.
- Radiant transfer of thermal energy from one place to another occurs by means of electromagnetic waves.
- When electromagnetic radiation falls on an object, it will be partially reflected, partially transmitted and partially absorbed.
- The rate of emission or absorption of radiant heat will depend upon the:
  - temperature difference between the object and the surrounding environment
  - surface area and surface characteristics of the object
  - wavelength of the radiation.

KEY QUESTIONS

1 Light is shone on an object.

a List three interactions that can occur between the light and the object.

b Which of the interactions from part (a) are associated with a rise in temperature?

2 The wavelength and frequency of emitted radiation depends on the internal energy of an object. Complete the sentences below by choosing the correct option from those provided in brackets.

The higher the temperature of the object, the [**higher**/**lower**] the frequency and the [**longer**/**shorter**] the wavelength of the radiation emitted. For example, if a particular object emits radiation in the visible range, a cooler object could emit light in the [**infrared**/**ultraviolet**] range of the electromagnetic spectrum.

3 Which of the following will affect the rate at which an object radiates thermal energy?

A its temperature

B its colour

C its surface nature (shiny or dull)

D none of these

E all of A–C

4 Why is it impossible for heat to travel from the Sun to the Earth by conduction or convection?

5 Thermal imaging technology can be used to locate people lost in the Australian bush. How can thermal imaging technology ‘see’ people when the human eye cannot?

6 Three identical, sealed beakers are filled with near-boiling water. One beaker is painted matte black, one is dull white and the third is gloss white.

a Which beaker will cool fastest?

b Which beaker will cool slowest?

7 Computer chips generate a lot of thermal energy that must be dispersed for the computer to function efficiently. Devices called heat sinks are used to help this process. Suggest what properties of a material would make it a good heat sink.

Chapter review

02

KEY TERMS

absorption	emission	insulator
conduction	emitted	internal energy
conductor	evaporation	passive heating
convection	frequency	radiation
electromagnetic spectrum	incident	wavelength

The following information applies to questions 1 and 2.

A household is considering changing from traditional incandescent lighting to LEDs. Using the respective percentage efficiencies from Table 2.1.1 on page 30, calculate:

1 The electrical energy effectively converted to light for each 1 kJ supplied for
 a an incandescent light
 b an LED.

2 How much more does an incandescent light cost to run than an LED light based on the efficiencies given in the table?

The following information applies to questions 3–6.

A student places a heating element and an electric whisk in an insulated container with 200 mL of water initially at 20.0°C, which is open at the top. He calculates that the heater adds 1850 J of heat energy to the water, and the whisk does 520 J of work on the water.

3 What is the total potential change in internal energy of the water?

4 If there was no heat loss from the container, calculate the final temperature the water would be expected to reach. Use $c_{water} = 4180\,J\,kg^{-1}\,K^{-1}$.

5 The student measured the final temperature and found it to only be 21.50°C due to losses to the air through the open top. How much of the energy supplied by the whisk and heater remained to increase the internal energy of the water?

6 Calculate the efficiency of the whisk and heater in heating the water once heat losses to the air are taken into account.

7 Two different objects are in thermal contact with each other. The objects are at different temperatures. What do the temperatures of the two objects determine?
 A The process by which thermal energy is transferred.
 B The heat capacity of each object.
 C The direction of transfer of thermal energy between the objects.
 D The amount of internal energy in each object.

8 Thermal energy may be transferred
 a in a fluid as a result of density changes of the fluid
 b in a non-metallic substance as a result of lattice vibrations.

Which process—conduction, convection or radiation—applies to each energy transfer?

9 Suppose you are sitting next to a fireplace in which there is a fire burning. One end of a metal poker has been left in the fire. Explain:
 a why you eventually feel the handle of the poker get hot
 b why you feel warm
 c how heat is lost from the room.

10 The interior of a thermos bottle is silvered to minimise heat transfer due to which one or more of the following processes: conduction, convection or radiation?

11 A house is fitted with an electric heater in one corner in order to heat a whole room. By which process or processes (conduction, convection or radiation) is the room heated when the heater is in operation?

12 The Sun continuously radiates energy into space, some of which reaches the Earth. The average temperature of the surface of the Earth remains about 300 K. Over the short-term why doesn't the average global temperature rise as the Sun's energy reaches it?

13 Explain the function of the evacuated enclosure between the walls of a vacuum flask.

14 Describe an experiment that would prove that a shiny, white surface is a poorer absorber of heat radiation than a dull or black surface. Explain one way this fact is used in everyday life.

15 A premature baby in an incubator can be dangerously cooled even when the air temperature in an incubator is warm. Explain why.

16 If you are lost in the snow, you are advised to build yourself a snow cave. In terms of thermal conductivity, explain how it is possible to stay warm inside a cave made of snow or ice.

The following information applies to questions 17 and 18.
An aluminium can with a paper label is left in a deep freeze for some time. On taking the can out of the freezer, your hand sticks to the cold aluminium but not the label.

17 Which, paper or can, will be at the lower temperature?

18 The label is peeled off the can. Which, paper or can, will return to normal room temperature first? Explain.

19 Design a system for a hot water storage tank that would permit hot water to be taken from tank and replacement cold water to enter without the two mixing. Include in your design appropriate connections to allow the water to be circulated without the need for pumps.

20 A day will feel warm if the air is at 24°C. However, a pool at 24°C will feel quite cool when swimming. Explain, using your understanding of heat processes, why this is the case.

CHAPTER 03

Particles in the nucleus

Radioactive emission is important in medicine for treating and detecting cancers, and for many imaging purposes. It also has many industrial applications, including power generation (covered in Chapter 4). This chapter examines the instability in the nucleus that leads to radioactive emissions. You will learn the properties of alpha, beta and gamma radiation, including balancing nuclear equations and predicting radiation types. You will also understand the importance of the rate of decay and half-lives of radioactive substances in determining appropriate nuclides for medical and industrial applications, how a radiation dose is measured, and the safe limits for doses of different types of radiation.

Science as a Human Endeavour

Qualitative and quantitative analyses of relative risk (including half-life, absorbed dose and dose equivalence) are used to inform community debates about the use of radioactive materials and nuclear reactions for a range of applications and purposes, including:

- radioisotopes are used as diagnostic tools and for tumour treatment in medicine.

Science Understanding

- the nuclear model of the atom describes the atom as consisting of an extremely small nucleus which contains most of the atom's mass, and is made up of positively charged protons and uncharged neutrons surrounded by negatively charged electrons
- nuclear stability is the result of the strong nuclear force which operates between nucleons over a very short distance and opposes the electrostatic repulsion between protons in the nucleus
- alpha and beta decay are examples of spontaneous transmutation reactions, while artificial transmutation is a managed process that changes one nuclide into another
- some nuclides are unstable and spontaneously decay, emitting alpha, beta (+/–) and/or gamma radiation over time until they become stable nuclides
- alpha, beta and gamma radiation have different natures, properties and effects
- each species of radionuclide has a half-life which indicates the rate of decay

  This includes applying the relationship

  $$N = N_0 \; \frac{1}{2}^{\,n}$$

- the measurement of absorbed dose and dose equivalence enables the analysis of health and environmental risks

  This includes applying the relationships

  absorbed dose $= \frac{E}{m}$, dose equivalent = absorbed dose × quality factor

3.1 Atoms, isotopes and radioisotopes

Many people mistakenly think that they never come into contact with radioactive materials or the radiation that these materials produce. However, the Earth is a radioactive planet and it is impossible to avoid exposure to radioactivity. Human senses cannot detect the radiation from radioactive atoms. Human beings are biologically adapted to cope with the background level of radiation; however, high-energy radiation in larger than normal doses can be damaging to living tissue. Radiation and radioactive elements can also be used in a variety of applications that are beneficial, like medicine. These radioactive atoms, or radioisotopes, will be discussed in this section.

ATOMS

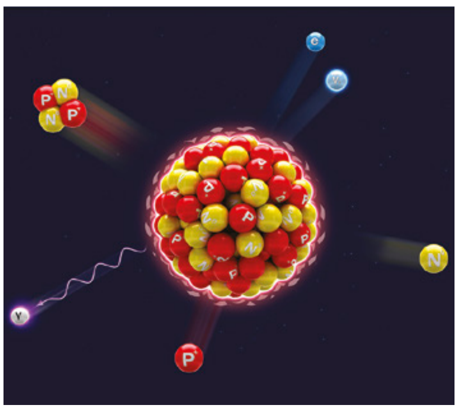


FIGURE 3.1.1 Radiation is spontaneously emitted from a radioactive nucleus.

If an atom is **radioactive**, it will spontaneously emit **radiation** from its nucleus. Figure 3.1.1 shows this radiation emitted in the form of particles and electromagnetic energy. To understand radiation and radioactivity, it is necessary to know about the structure of the atom. The central part of an atom, the **nucleus**, consists of particles known as **protons** and **neutrons**. Collectively, these particles are called **nucleons** and are almost identical in mass and size.

The nucleons have very different electrical properties. Protons have a positive charge and neutrons are electrically **neutral** so they have no charge. The nucleus contains nearly all the atom's mass. Most of the atom is empty space, occupied only by negatively charged particles called **electrons**, which surround the nucleus. These are much smaller and lighter than protons or neutrons. The diameter of the nucleus is about ten thousand to one hundred thousand times smaller than that of the atom. Almost the entire mass of an atom is concentrated in its nucleus; the total mass of its electrons is less than one thousandth of the mass of the atom.

Figure 3.1.2 shows the structure of a typical atom, with an expanded view of the nucleus. Radioactive decay and nuclear reactions result from interactions within or between nuclei.

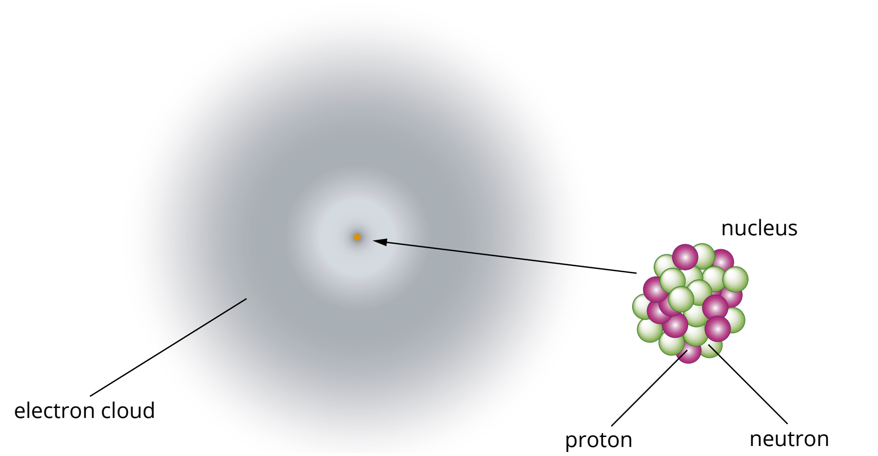


FIGURE 3.1.2 The nucleus of an atom occupies about 10^{-12} of the volume of the atom, yet it contains more than 99% of its mass. Atoms are mostly empty space. (Note, this atom is not drawn to scale.)

A particular atom can be identified by using atomic symbols that have the format shown in Figure 3.1.3.

mass number

$${}^{A}_{Z}\mathrm{X}$$ element symbol

atomic number

FIGURE 3.1.3 Atomic notation.

The **mass number** (A), sometimes called the atomic mass, is the total number of protons and neutrons in the nucleus.

The **atomic number** (Z) is the number of protons in the nucleus.

The number of neutrons (N) is given by $N = A - Z$.

Atoms with the same number of protons belong to the same element. For example, if an atom has six protons in its nucleus (i.e. $Z = 6$), then the atom must be carbon. The number of neutrons does not affect which element the atom is, but it does affect the mass of the atom. Figure 3.1.4 shows how the size of the nucleus depends on the mass number. The more protons and neutrons there are in a nucleus, the heavier and larger it is.

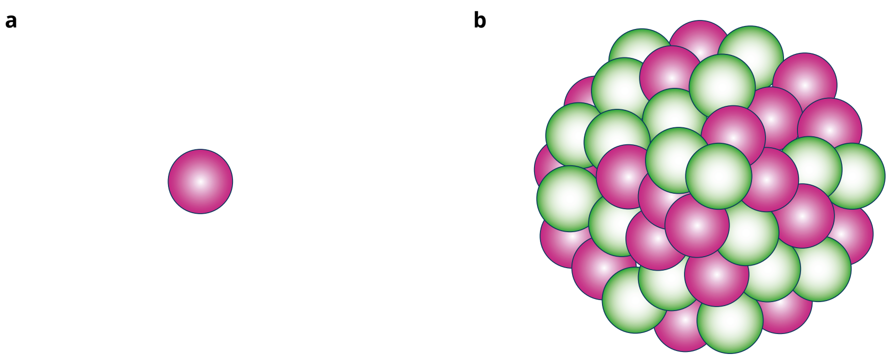


FIGURE 3.1.4 (a) and (b) are both nuclei, however the hydrogen atom (a) has a very different size nucleus to a uranium atom (b). (Note, the nuclei are not drawn to scale.)

In an electrically neutral atom the number of electrons is equal to the number of protons. For example, any neutral atom of uranium ($Z = 92$) has 92 protons in the nucleus and 92 electrons in the electron cloud.

Worked example 3.1.1

WORKING WITH ATOMIC NOTATION

How many protons, neutrons, nucleons and electrons are there in $^{197}_{79}Au$?	
Thinking	**Working**
The lower number is the atomic number, Z. This is the number of protons.	atomic number, $Z = 79$ This nuclide has 79 protons.
The upper number is the mass number, A. This indicates the number of particles in the nucleus, i.e. the number of nucleons.	mass number, $A = 197$ This nuclide has 197 nucleons.
The number of neutrons, N, is the difference between the mass number, A (the number of nucleons), and the atomic number, Z (the number of protons).	$N = A - Z$ $= 197 - 79$ $= 118$ This nuclide has 118 neutrons.
In an electrically neutral atom the number of protons = the number of electrons.	The nuclide has 79 protons, so the atom will have 79 electrons in the electron cloud.

Worked example: Try yourself 3.1.1

WORKING WITH ATOMIC NOTATION

How many protons, neutrons, nucleons and electrons are there in $^{252}_{92}U$?

PHYSICSFILE

Neutron stars

In the universe there are objects whose density is almost equal to that of nuclear matter. These are called neutron stars. They are gigantic balls with a radius of ten or more kilometres, and are made only of neutrons—something like a gigantic (protonless) atomic nucleus. If a one-litre juice carton was filled up with this type of matter it would weigh a thousand times more than the largest Egyptian pyramid.

PHYSICSFILE

Mass Number and Atomic Mass

Mass number referred to in this text is the number of protons plus neutrons. An element can have a number of stable isotopes, for example although Chlorine has many possible isotopes there are two commonly found stable isotopes found in the following ratios, 24.2% Chlorine-37 and 75.8% Chlorine -35. Chlorine-35 has a mass number of 35 and chlorine-37 a mass number of 35.

However the atomic mass you will find on a periodic table is 35.45 which is the weighted average of the commonly found isotopes.

PHYSICSFILE

Heavy water

A compound of oxygen and deuterium has identical chemical properties to ordinary water. However, the molecular mass of ordinary water is about 18 (16 + 1 + 1) while the molecular mass of water containing deuterium is 20 (16 + 2 + 2). Thus, water that contains deuterium has a higher density (by about 11%) and is commonly known as 'heavy water'.

ISOTOPES

All atoms of a particular element will have the same number of protons, but may have a different number of neutrons. For example, lithium exists naturally in two different forms. One form has three protons and three neutrons. The other has three protons and four neutrons. These different forms of lithium are called **isotopes** of lithium. These isotopes are illustrated in Figure 3.1.5.

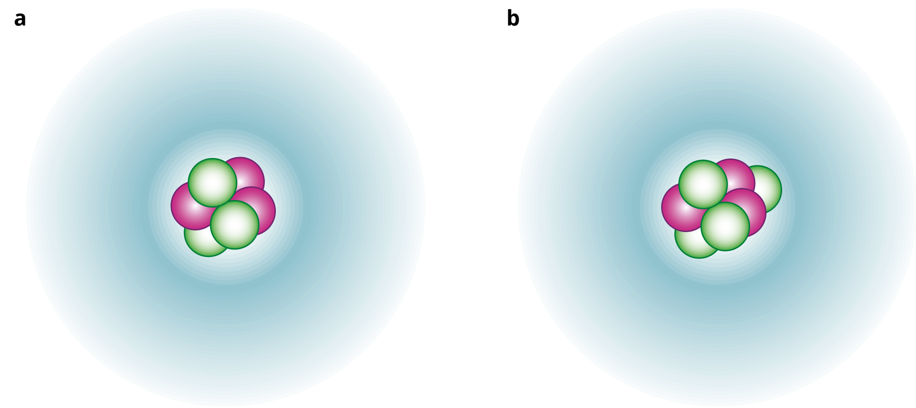


FIGURE 3.1.5 Two different isotopes of lithium: (a) ${}^{6}_{3}Li$ and (b) ${}^{7}_{3}Li$.

Isotopes are atoms that have the same number of protons but different numbers of neutrons. Isotopes have the same chemical properties but different physical properties such as density and volume.

The term **nuclide** is used when referring to a particular nucleus. For example, lithium-6 is a nuclide which has three protons and three neutrons.

There are three isotopes of hydrogen: the nuclide with one proton is called hydrogen, the nuclide with one proton and one neutron is called **deuterium**, and the nuclide with one proton and two neutrons is called **tritium**.

Worked example 3.1.2

WORKING WITH ISOTOPES

Consider the isotope of molybdenum, ${}^{95}_{42}Mo$. Work out the number of protons, nucleons and neutrons in this isotope.	
Thinking	**Working**
The lower number is the atomic number, Z. This is the number of protons.	atomic number, $Z = 42$ This nuclide has 42 protons.
The upper number is the mass number, A. This indicates the number of particles in the nucleus, i.e. the number of nucleons.	mass number, $A = 95$ This nuclide has 95 nucleons.
Subtract the atomic number, Z, from the mass number, A, to find the number of neutrons, N.	$N = A - Z$ $= 95 - 42$ $= 53$ This isotope has 53 neutrons.

Worked example: Try yourself 3.1.2

WORKING WITH ISOTOPES

Consider the isotope of thorium, ${}^{230}_{90}Th$. Work out the number of protons, nucleons and neutrons in this isotope.

RADIOISOTOPES

Most of the atoms that make up the world around us are stable. Their nuclei have not altered in the billions of years since they were formed. These atoms will stay unchanged in the future. There are about 270 stable isotopes in nature. Tin ($Z = 50$) has ten stable isotopes, while aluminium ($Z = 13$) has just one.

There are also naturally occurring isotopes that are unstable. In a process known as **spontaneous transmutation** an unstable nucleus may become more stable by spontaneously emitting a particle and so change into a different element or isotope. Unstable atoms are **radioactive** and an individual radioactive isotope is known as a **radioisotope**. For example, carbon has two stable isotopes: carbon-12 and carbon-13. Carbon also has one naturally occurring isotope that is unstable: carbon-14. The nucleus of a carbon-14 atom may spontaneously decay into a different substance, emitting high-energy particles. A known radioactive substance is identified by the radiation warning symbol shown in Figure 3.1.6.

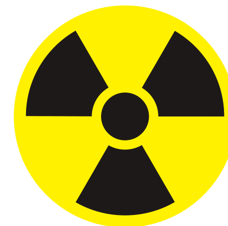


FIGURE 3.1.6 This symbol is used to label and identify a radioactive source.

The periodic table is shown in Figure 3.1.7. Every isotope of every element with an atomic number greater than that of bismuth ($Z = 83$) is radioactive, as is technetium 99m ($Z = 43$) which is artificially produced. The other first 92 elements are naturally occurring.

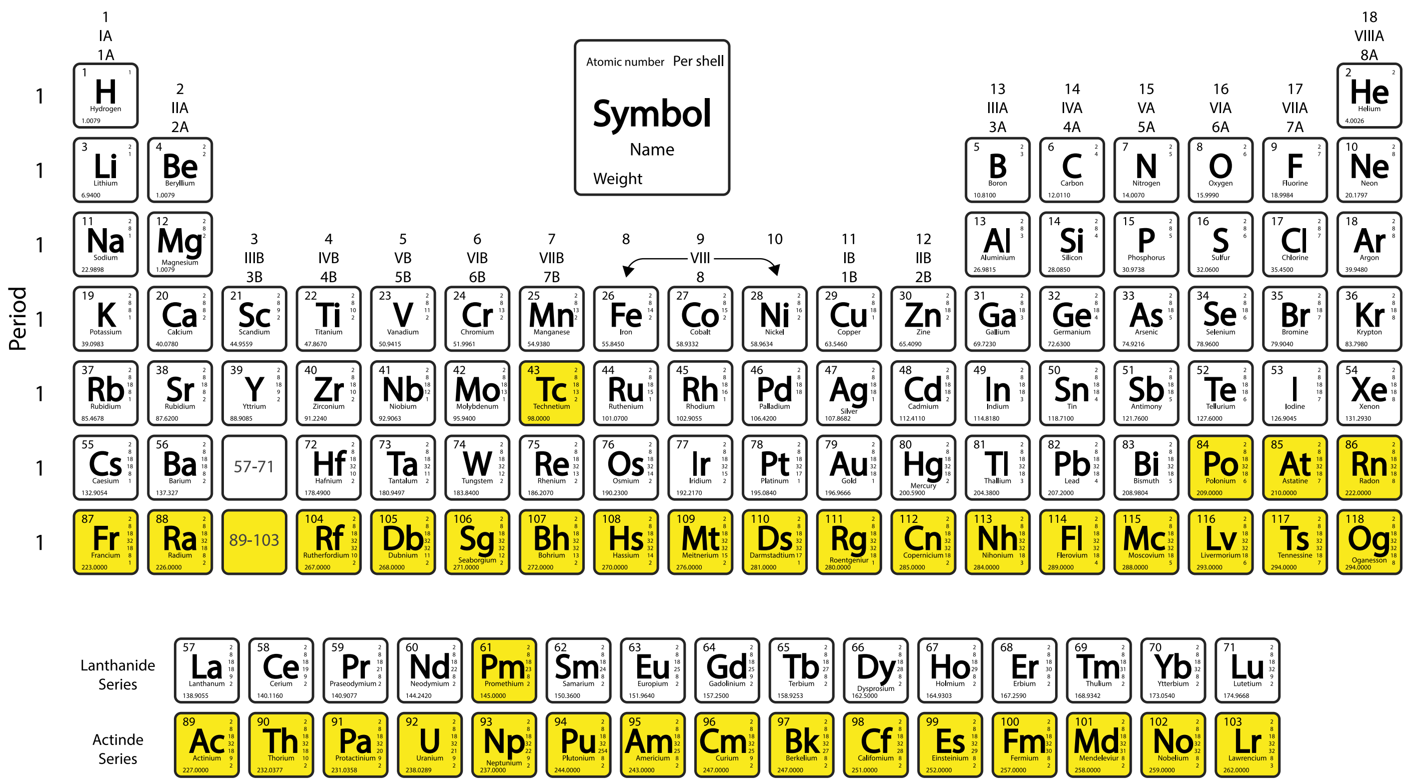


FIGURE 3.1.7 The periodic table of the elements. All isotopes of the elements shaded in yellow are radioactive.

Most of the elements found on Earth have naturally occurring radioisotopes; there are about 200 of these natural radioisotopes. During the twentieth century, an enormous number of radioisotopes were also artificially produced in a process known as **artificial transmutation**. This is a managed process that changes the number of protons, thus changing the identity of the nuclide. Artificial radioisotopes are produced in nuclear reactors or particle accelerators. Most of the radioisotopes used in industry, medicine and for scientific research are artificially produced.

ARTIFICIAL TRANSMUTATION: HOW RADIOISOTOPES ARE MANUFACTURED

Radioisotopes have many medical and industrial applications, such as for the diagnosis of cancer. Most of the radioisotopes that are used in these applications are synthesised by artificial transmutation. There are now about 3000 different artificial radioisotopes. In the periodic table, every element with an atomic number greater than 92 (i.e. that appears after uranium) is radioactive and is produced artificially. The periodic table in Figure 3.1.7 includes recently discovered elements nihonium (Nh), moscovium (Mc), tennessine (Ts) and oganesson (Og), which were recognised formally by IUPAC, the International Union for Pure and Applied Chemistry, in 2016. You may see some older versions of the periodic table that have the elements' temporary names of ununtrium (Uut), ununpentium (Uup), ununseptium (Uus) and ununoctium (Uuo), given before they were officially recognized by IUPAC. These names come from the Latin for their atomic numbers.

One of the ways that artificial radioisotopes are manufactured is by neutron absorption. In this method, a sample of a stable isotope is placed inside a nuclear reactor and bombarded with neutrons. For example, artificial radioisotopes for medical and industrial uses are manufactured in the core of the Lucas Heights reactor at the ANSTO (Australian Nuclear Science and Technology Organisation) facility in Sydney (Figure 3.1.8). This is Australia's only nuclear reactor facility and has been operating since 1958. The original reactor was replaced by the OPAL (Open Pool Australian Light-water) reactor in 2007. Nuclear reactors will be discussed in Chapter 4.

FIGURE 3.1.8 The nuclear reactor at ANSTO, Lucas Heights, Sydney.

When one of the bombarding, or irradiating, neutrons collides with a nucleus of the stable isotope, the neutron is absorbed into the nucleus. This may create an unstable isotope of the sample element.

One of the most widely used radioisotopes is cobalt-60. It does not exist in nature, but is artificially produced in the core of a nuclear reactor by bombarding stable cobalt-59 with neutrons. Figure 3.1.9 shows a large cobalt-59 nucleus with a small neutron speeding towards it. The arrow indicates that the product of the collision is a radioactive cobalt-60 nucleus.

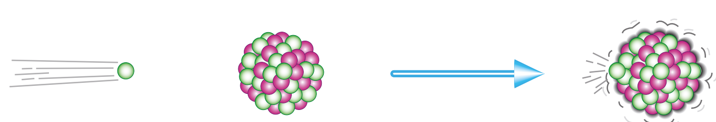


FIGURE 3.1.9 A neutron colliding with ^{59}Co nuclei to form ^{60}Co.

PHYSICSFILE

Oganesson (Og)

The element with the highest atomic number and highest atomic mass so far discovered is oganesson (Og), originally named as unumoctium (Uuo). Three atoms of this element were made in a particle accelerator in 2006 when calcium-48 nuclei were bombarded with californium-249 nuclei. The 20 protons of calcium combined with the 98 of californium to make just one or two atoms of Og. Og is very unstable and decays very rapidly, with a half-life of less than 1 ms. You will learn more about half-lives in Section 3.2.

STABILITY OF NUCLEI AND WHY RADIOACTIVE NUCLEI ARE UNSTABLE

In Figure 3.1.10, stable and radioactive isotopes have been plotted according to their number of protons (atomic number, Z) and their number of neutrons (N). The stable nuclides that exist in nature are indicated by purple squares. The radioisotopes that emit radiation are indicated by the black plus and minus signs and α symbols, and will be discussed in the next section.

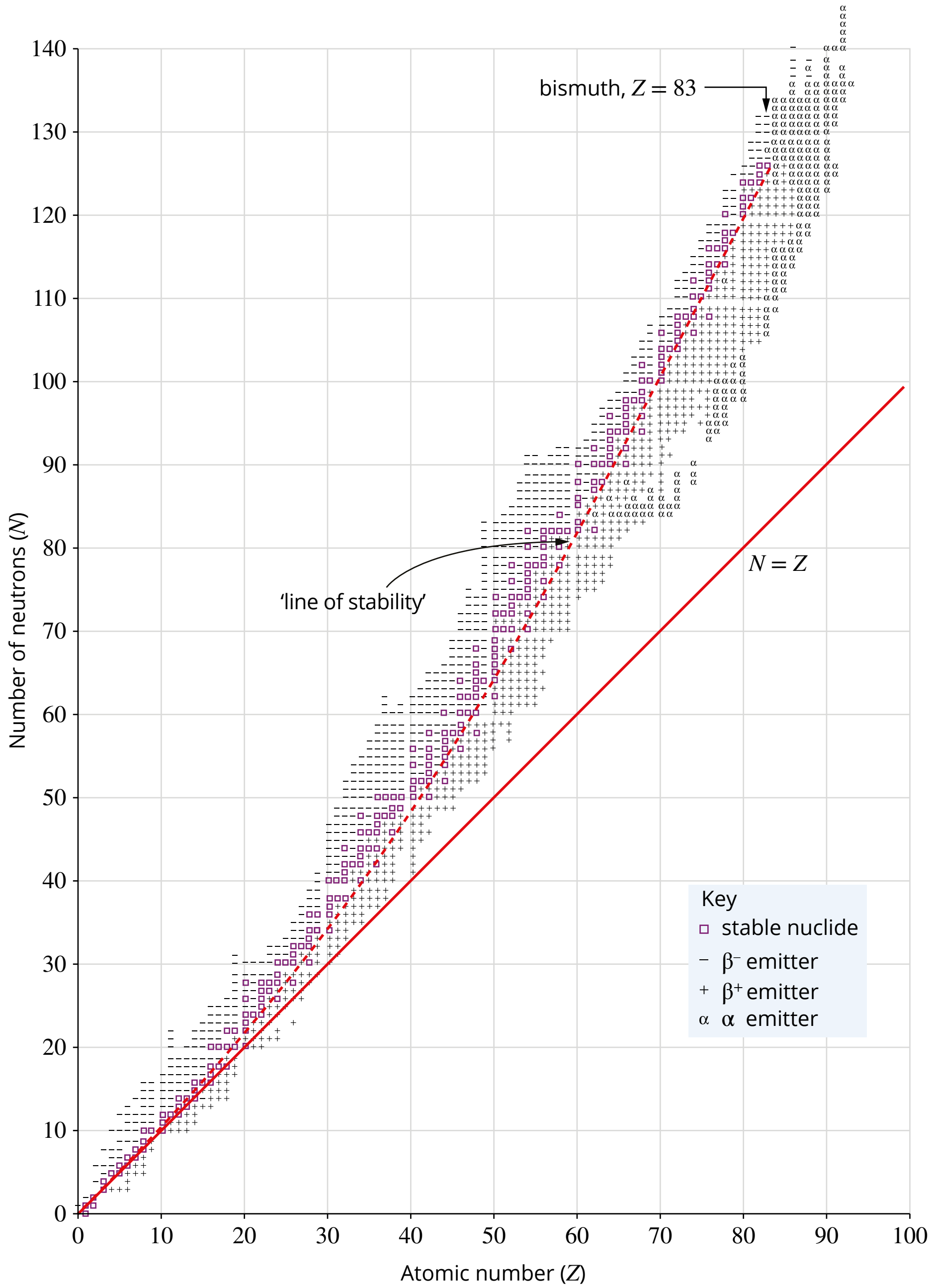


FIGURE 3.1.10 A chart showing stable and radioactive isotopes, plotted according to their number of protons (atomic number) and number of neutrons.

Within the nucleus, protons are very close to other protons. This should seem odd as protons exert strong **electrostatic forces** of repulsion on each other. Electrostatic forces act between charged particles and can act over relatively large distances. In the nucleus, this means that each proton strongly repels every other proton so this force is trying to make the nucleus break apart. As it doesn't, there must be other factors to consider. A force known as the **strong nuclear force** is also acting.

The strong nuclear force is a force of attraction that holds the nucleus together and acts between every nucleon regardless of their charge. This force acts like 'nuclear glue'. Neutrons are attracted to nearby neutrons and protons. Protons are also attracted to nearby neutrons and protons. However, this force only acts over relatively short distances so for nucleons on the opposite sides of a large nucleus, this force is not significant.

In a stable nucleus, there is a delicate balance between the repulsive electric force and the attractive strong nuclear force. For example, bismuth-209, the heaviest stable isotope, has 83 protons and 126 neutrons. Here, the electrostatic repulsion of the protons is balanced by the strong attractive nuclear forces between the nucleons to make the nucleus stable. Compare this with bismuth-211. Its two extra neutrons upset the balance between forces. The nucleus of ^{211}Bi is unstable and it ejects an alpha particle (a charged $^{4}_{2}He$ nucleus) in an attempt to become more stable.

From Figure 3.1.10 it is evident that there is a 'line of stability' (indicated by the curved red dashed line on the graph) along which the stable nuclei tend to cluster. Nuclei away from this line are unstable.

For small nuclei with atomic numbers up to about 20, the ratio of neutrons to protons in stable nuclei is close to one. However, as the nuclei become bigger, this ratio increases for stable nuclei. Zirconium (Z = 40) has a neutron-to-proton ratio of about 1.25, while for mercury (Z = 80) the ratio is close to 1.66. This indicates that for higher numbers of protons, nuclei must have even more neutrons to remain stable. These neutrons act to dilute the repelling forces that exist between the extra protons.

Elements with more protons than bismuth (Z = 83) simply have too many repulsive charges in the nucleus. Additional neutrons are unable to stabilise these nuclei. All of these elements are unstable and radioactive. Figure 3.1.11 illustrates stable and unstable nuclei.

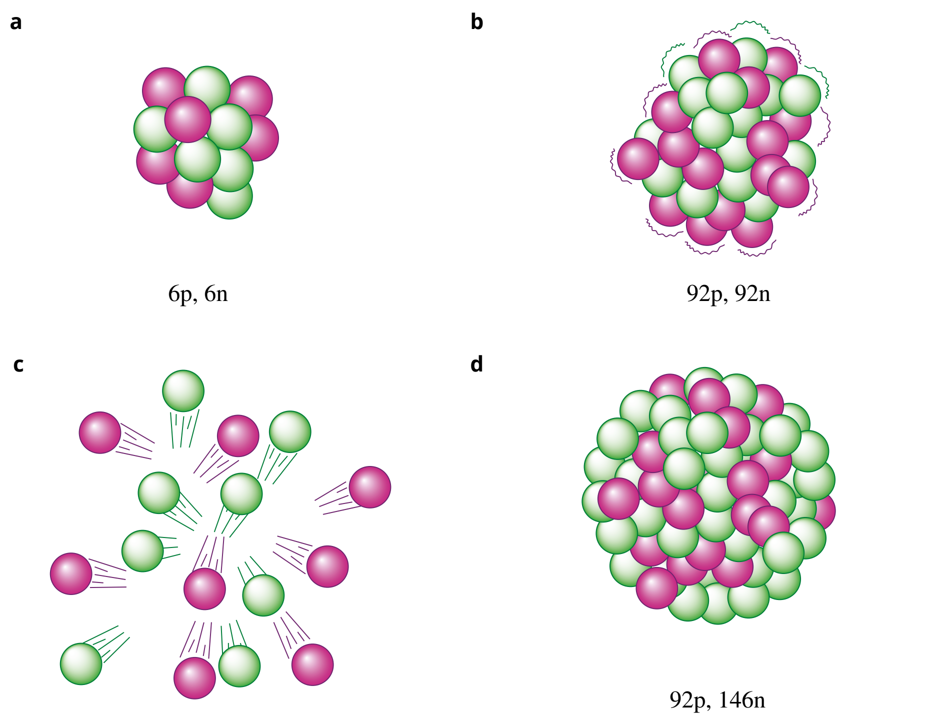


FIGURE 3.1.11 Stable and unstable nuclei. (a) A small nucleus such as carbon-12 is stable. This is because the electrostatic force of repulsion that acts between the protons is overcome by the strong nuclear force of attraction. (b) and (c) A large nucleus with equal numbers of protons and neutrons cannot exist. The electrostatic forces of repulsion between the protons would overcome the strong nuclear forces. (d) Additional neutrons increase the stability of large nuclei. The extra neutrons increase the influence of the strong nuclear force and act like a 'nuclear glue', holding the nucleus together.

EXTENSION

Quarks and other subatomic particles

The understanding of the atom has changed greatly in the past 100 years. It was once thought that atoms were like miniature billiard balls: solid and indivisible. The word 'atom' comes from the Greek *atomos*, meaning indivisible. That idea was changed forever when the first subatomic particles—the electron, the proton and then the neutron—were discovered between 1897 and 1932.

Since World War II, further research has uncovered about 300 other subatomic particles. These include pi-mesons, mu-mesons, kaons, tau leptons and neutrinos.

Figure 3.1.12 shows a photograph taken at CERN (the European Organization for Nuclear Research) in Switzerland. It shows hundreds of charged subatomic particles spilling out from the collision of a high-energy oxygen nucleus with a lead nucleus in a lead target. Figure 3.1.13 shows an underground tunnel with a blue tube stretching into the distance. It is part of the particle accelerator at CERN. It accelerates protons from rest to 99.99995% of the speed of light in under 20 seconds.

For many years, physicists found it difficult to make sense of this array of subatomic particles. It was known that one family of particles called the leptons had six members: electron, electron-neutrino, muon, muon-neutrino, tau and tau-neutrino.

Then, in 1964, Murray Gell-Mann proposed a simple theory. He suggested that most subatomic particles were composed of a number of more fundamental particles called quarks. Currently, it is accepted that there are six different quarks: up, down, charmed, strange, top and bottom. The latest quark to be identified was the top quark, whose existence was confirmed in 1995. A proton consists of two up quarks and one down quark, and a neutron consists of one up quark and two down quarks. Subatomic particles that consist of quarks are known as hadrons. Leptons, such as electrons, are indivisible point particles; they are not composed of quarks.

A significant amount of effort and money has been directed to testing Gell-Mann's theory, both theoretically and experimentally. This has involved the construction of larger and larger particle accelerators such as Fermilab in Chicago and CERN in Geneva. Figure 3.1.14 shows particle tracks from a proton–proton collision seen by the Large Hadron Collider detector at CERN. CERN was the site of the discovery of the Higgs boson in 2012, a fundamental particle needed to explain the property of mass. Australia built its own particle accelerator—a synchrotron—next to Monash University in Victoria. This began operating in 2007, and is still used in many types of research today.

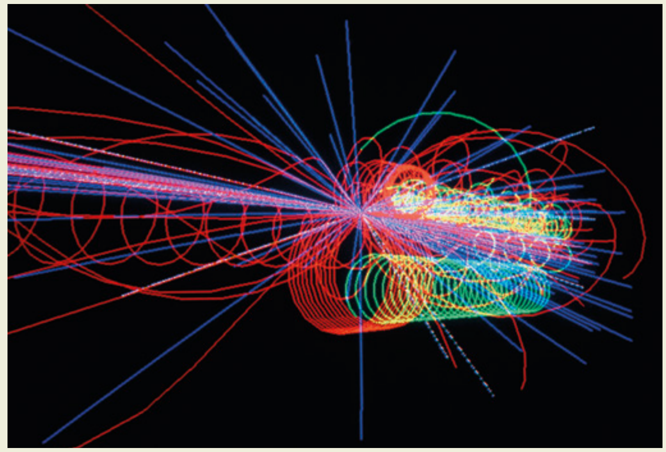


FIGURE 3.1.12 Collision of subatomic particles at CERN (the European Organization for Nuclear Research) in Switzerland.

FIGURE 3.1.13 Particle accelerator at CERN in Switzerland.

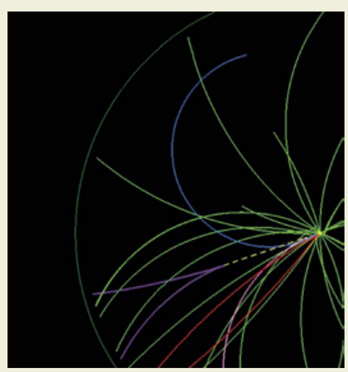


FIGURE 3.1.14 Collision between protons in the Large Hadron Collider at CERN.

While the current theory suggests that quarks and leptons are the ultimate fundamental particles that make up matter, and cannot be further divided, scientific theories and models can change as new experimental data are obtained. Are quarks and leptons made of smaller particles again? Time will tell.

3.1 Review

SUMMARY

- The nucleus of an atom consists of positively charged protons and neutral neutrons. Collectively, protons and neutrons are known as nucleons. Negatively charged electrons surround the nucleus.
- The nucleus of the atom is extremely small but contains most of the atom's mass.
- The atomic number, *Z*, is the number of protons in the nucleus. The mass number, *A*, is the number of nucleons in the nucleus, i.e. the combined number of protons and neutrons. Elements are represented as $^{A}_{Z}X$.The number of neutrons, $N = A - Z$.
- Isotopes of an element have the same number of protons but different numbers of neutrons. Isotopes of an element are chemically identical to each other, but have different physical properties.
- An unstable isotope—a radioisotope—may spontaneously decay by emitting a particle from the nucleus. This is called spontaneous transmutation if the number of protons, and hence the identity of the element, is changed.
- Artificial isotopes are produced by a process called artificial transmutation, which changes the number of protons, thus changing the identity of the nuclide. This commonly takes place as a result of neutron bombardment in the core of a nuclear reactor.
- Nuclear stability is the result of the strong nuclear force between nucleons over a very short distance, which opposes electrostatic repulsion between protons.

KEY QUESTIONS

1 What is the collective term used for protons and neutrons?

2 How many protons and how many neutrons are in the $^{197}_{79}Au$ nuclide?

3 How many nucleons are there in the $^{235}_{92}U$ nuclide?

4 Determine the number of protons, neutrons and nucleons in the following nuclides. You may need to refer to the periodic table in Figure 3.1.7 on page 49.
- **a** chlorine-35
- **b** plutonium-239

5 Which one or more of the following nuclides have seven neutrons in the nucleus? You may need to refer to the periodic table in Figure 3.1.7.
- **A** carbon-12
- **B** carbon-13
- **C** carbon-14
- **D** nitrogen-14

6 How is the number of electrons in a neutral atom found from information given in the periodic table?

7 Explain the meaning of the term isotope.

8 Krypton-84 is stable, but krypton-89 is radioactive. Imagine that you have just one atom of each isotope.
- **a** Are their atomic numbers and mass numbers the same or different? Justify your answers.
- **b** Compare the way these atoms would interact chemically with other atoms.

9 What is the difference between a stable isotope and a radioisotope?

10 Can a natural isotope be radioactive? If so, give an example of such an isotope.

11 Why is the number of neutrons greater than the number of protons in a nucleus as the elements get heavier?

3.2 Radioactivity

Around the beginning of the twentieth century, scientists such as Marie Curie, pictured in Figure 3.2.1, were investigating the newly discovered radioactive substances polonium and radium. Ernest Rutherford and Paul Villard found that there were three different types of emission from these mysterious substances. They named them alpha, beta and gamma radiation.

FIGURE 3.2.1 Marie Curie, pioneer of research into radioactivity.

Further experiments showed that the alpha and beta emissions were actually particles expelled from the nucleus. Gamma radiation was found to be high-energy **electromagnetic radiation** (similar to visible light but of significantly higher energy) expelled from the nucleus. The term radioactive decay refers to the process that emits these particles and radiation from a nucleus.

The origin and nature of these radiations will be discussed in this section.

ALPHA (α) DECAY

When a heavy unstable nucleus undergoes radioactive decay, it may eject an **alpha particle**. This is a positively charged particle that consists of two protons and two neutrons, with no orbiting electrons. An alpha particle, symbol α, is identical to a helium nucleus and can also be written as ${}^{4}_{2}\mathrm{He}$ or ${}^{4}_{2}\mathrm{He}^{2+}$.

Uranium-238 is radioactive and may decay by emitting an alpha particle from its nucleus. Figure 3.2.2 shows the unstable nucleus of uranium-238 ejecting an alpha particle. This can be represented in a nuclear equation, which shows the changes occurring in the nuclei. Electrons are not considered in these equations, only the nucleons. The equation for the alpha decay of uranium-238 is:

$$ {}^{238}_{92}\mathrm{U} \rightarrow {}^{234}_{90}\mathrm{Th} + {}^{4}_{2}\mathrm{He} + \text{energy} $$

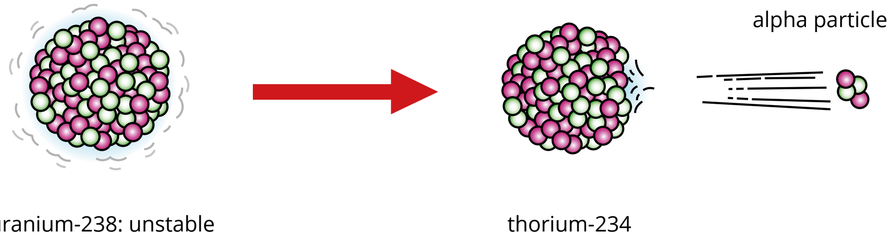


FIGURE 3.2.2 Alpha emission from uranium-238.

The **parent nucleus** ${}^{238}_{92}\mathrm{U}$ has spontaneously emitted an alpha particle (α) and has changed into a completely different element, ${}^{234}_{90}\mathrm{Th}$. Thorium-234 is called the **daughter nucleus**. The energy released is mostly kinetic energy carried by the fast-moving alpha particle.

When an atom changes into a different element, it is said to undergo a **nuclear transmutation**. In nuclear transmutations, electric charge is conserved. This results in the conservation of atomic number, i.e. the number of protons. The sums of atomic numbers on both sides of a nuclear equation must be equal. In the uranium decay equation, the atomic number, or the number of protons, is conserved: 92 = 90 + 2. The mass number is also conserved: 238 = 234 + 4.

In any nuclear reaction, including radioactive decay, atomic and mass numbers are conserved. Energy is released during these decays.

PHYSICSFILE

Radioactive lamps

The wicks or mantles used in old-style camping lamps are slightly radioactive. They contain a radioisotope of thorium, an alpha-particle emitter. They have not been banned from sale so far because they contain only small amounts of the radioisotope and can be used safely by taking simple precautions such as washing hands and avoiding inhalation or ingestion.

However, a scientist from the Australian National University in Canberra has called for these mantles to be banned because they tend to crumble and turn to dust as they age. If this dust were inhaled, alpha particles could settle in someone's lung tissue, possibly causing cancers to form.

FIGURE 3.2.3 An old radioactive gas-light mantle.

Worked example 3.2.1

ALPHA DECAY

A polonium-212 nucleus is known to decay to a new element through the emission of an alpha particle. Determine the new element, write its symbol and write the decay equation.

Thinking	Working
From the periodic table, polonium-212 has 84 protons. Therefore its atomic number, *Z*, is 84 and its mass number, *A*, is 212.	It can be written $^{212}_{84}\text{Po}$.
The initial nucleus is $^{212}_{84}\text{Po}$ and is written on the left-hand side of the equation. The unknown nucleus is a result of alpha decay $\left(^{4}_{2}\text{He}\right)$ and is written on the right-hand side along with the alpha particle.	$^{212}_{84}\text{Po} \rightarrow {}^{A}_{Z}X + \alpha$ or $^{212}_{84}\text{Po} \rightarrow {}^{A}_{Z}X + {}^{4}_{2}\text{He}$
Charge must be conserved, so the total number of protons, *Z*, must be the same.	$84 = Z + 2$ $Z = 82$
The number of protons and neutrons, *A*, must also be the same.	$212 = A + 4$ $A = 208$
For the new element, $Z = 82$. From the periodic table, this is lead.	$^{208}_{82}\text{Pb}$
The decay equation can now be written.	$^{212}_{84}\text{Po} \rightarrow {}^{208}_{82}\text{Pb} + {}^{4}_{2}\text{He}$

Worked example: Try yourself 3.2.1

ALPHA DECAY

A radium-224 nucleus is known to decay to a new element through the emission of an alpha particle. Determine the new element, write the appropriate symbol and the decay equation.

BETA (β) DECAY

Many radioactive materials emit **beta particles**. There are two different types of beta particles: beta minus (β^-) and beta plus (β^+).

Beta minus (β^-)

This type of beta decay occurs when an electron is emitted from the nucleus of a radioactive atom, rather than from the electron cloud. This type of beta particle can be written as $^{0}_{-1}\beta$.

The atomic number of −1 indicates that the beta particle (the electron) has a single negative charge. The mass number of zero indicates that its mass is far less than that of a proton or a neutron.

Typically, beta-minus decay occurs if a nucleus has too many neutrons to be stable. A neutron spontaneously changes into a proton, a beta-minus particle (β^-, an electron), and an uncharged massless antimatter particle called an **antineutrino** ($\bar{\nu}$). This makes the nucleus more stable.

An example of an isotope that undergoes beta-minus decay is carbon-14. The other isotopes of carbon, carbon-12 and carbon-13, are both stable. Carbon-14 is unstable. It has too many neutrons and undergoes a beta-minus decay to become stable. One of the neutrons changes into a proton, emitting a beta-minus particle and an antineutrino in the process. Nitrogen-14 is then formed and energy is released. The beta-minus decay of carbon-14 is shown in Figure 3.2.4.

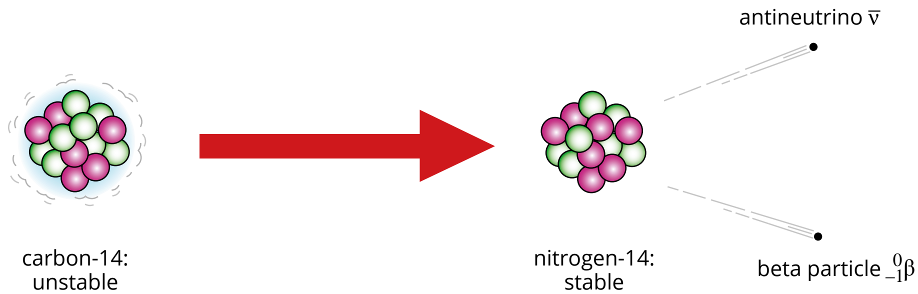


FIGURE 3.2.4 The beta-minus decay of carbon-14.

The nuclear equation for this decay is:

$${}^{14}_{6}\mathrm{C} \rightarrow {}^{14}_{7}\mathrm{N} + {}^{0}_{-1}\beta + \bar{\nu} + \text{energy}$$

The transformation taking place inside the nucleus is:

$${}^{1}_{0}\mathrm{n} \rightarrow {}^{1}_{1}\mathrm{p} + {}^{0}_{-1}\beta + \bar{\nu}$$

Notice that, in all these equations, the atomic and mass numbers are conserved. The antineutrino has no charge and has so little mass that both its atomic and mass numbers are zero.

Worked example 3.2.2

BETA MINUS DECAY

A bismuth-212 nucleus is known to decay to a new element through the emission of a beta particle. Determine the new element, write the appropriate symbol and the decay equation.	
Thinking	**Working**
From the periodic table, bismuth-212 has 83 protons. Therefore its atomic number, *Z*, is 83 and its mass number, *A*, is 212.	It can be written ${}^{212}_{83}\mathrm{Bi}$.
The initial nucleus is ${}^{212}_{83}\mathrm{Bi}$ and is written on the left-hand side of the equation. The unknown nucleus is a result of beta-minus decay and is written on the right-hand side along with the beta-minus particle and an antineutrino.	${}^{212}_{83}\mathrm{Bi} \rightarrow {}^{A}_{Z}X + {}^{0}_{-1}\beta + \bar{\nu}$.
Charge must be conserved, so the total number of protons, *Z*, must be the same.	$83 = Z - 1$ $Z = 84$.
The number of protons and neutrons, *A*, must also be the same.	$212 = A + 0$ $A = 212$
For the new element, *Z* = 84. From the periodic table, this is polonium.	${}^{212}_{84}\mathrm{Po}$
The decay equation can now be written.	${}^{212}_{83}\mathrm{Bi} \rightarrow {}^{212}_{84}\mathrm{Po} + {}^{0}_{-1}\beta + \bar{\nu}$

Worked example: Try yourself 3.2.2

BETA MINUS DECAY

An astatine-219 nucleus is known to decay to a new element through the emission of a beta-minus particle. Determine the new element, write its symbol and write the decay equation.

Beta plus (β^+)

A different form of beta decay occurs when a nucleus has too many protons. In this case, a proton may spontaneously change into a neutron and emit a neutrino (ν) and a positively charged beta particle. This process is known as β^+ (beta-positive) decay. The positively charged beta particle is called a **positron**. Positrons (${}^{0}_{+1}\beta$) have the same properties as electrons, but their electrical charge is positive rather than negative.

An example of beta-plus decay is given by:

$${}^{12}_{7}\text{N} \rightarrow {}^{12}_{6}\text{C} + {}^{0}_{+1}\beta + \nu$$

The transformation taking place inside the nucleus is:

$${}^{1}_{1}\text{p} \rightarrow {}^{1}_{0}\text{n} + {}^{0}_{+1}\beta + \nu$$

Positrons are an example of antimatter. Other types of antimatter are antiprotons, antineutrons and antineutrinos. Antimatter exists throughout the universe, but the overwhelming majority of the universe is composed of matter.

GAMMA (γ) DECAY

After a radioisotope has emitted an alpha or beta particle the daughter nucleus usually has excess energy. The protons and neutrons in the daughter nucleus then rearrange slightly and offload this excess energy by releasing a **gamma ray**, ${}^{0}_{0}\gamma$.

Gamma rays are high-energy electromagnetic radiation and so have no mass, are uncharged and travel at the speed of light (3.0×10^8 m s^{-1}).

A common example of a gamma-ray emitter is iodine-131. It decays by beta and gamma emission to form xenon-131, as shown in Figure 3.2.5. The xenon nuclide is in an excited state, as indicated by the asterisk, and loses energy by gamma emission.

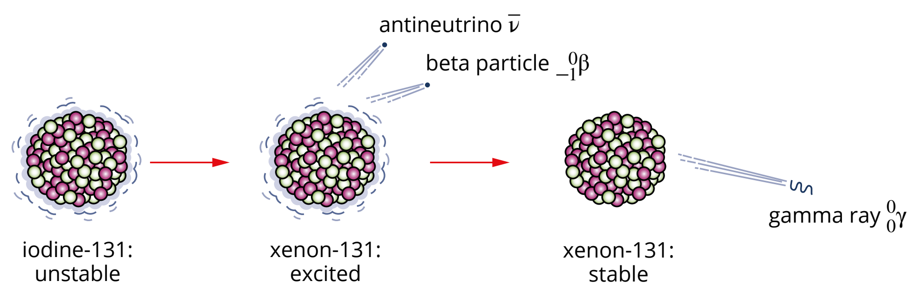


FIGURE 3.2.5 The gamma and beta decay of iodine-131.

The equations for this decay are:

$${}^{131}_{53}\text{I} \rightarrow {}^{131}_{54}\text{Xe}^* + {}^{0}_{-1}\beta$$

$${}^{131}_{54}\text{Xe}^* \rightarrow {}^{131}_{54}\text{Xe} + {}^{0}_{0}\gamma$$

Since gamma rays carry no charge and have no mass, they have no effect when balancing the atomic or mass numbers in a nuclear equation.

Worked example 3.2.3

RADIOACTIVE DECAY

An excited helium-3 atom decays by radioactive emission to form helium-3. The equation is: ${}^{3}_{2}\text{He}^{*} \rightarrow {}^{3}_{2}\text{He} + X$. Determine the atomic and mass numbers for *X* and identify the type of radiation being emitted.	
Thinking	**Working**
Balance the mass numbers.	The mass numbers of 3 are already balanced, so the mass number of *X* is zero.
Balance the atomic numbers.	The atomic numbers of 2 are already balanced, so the atomic number of *X* is zero.
X has an atomic number of zero and a mass number of zero.	*X* is a gamma ray, γ.

Worked example: Try yourself 3.2.3

RADIOACTIVE DECAY

After beta-minus decay from boron to carbon-12, the carbon-12 atom is in an excited state and decays further to a more stable form of carbon-12. The equation is ${}^{12}_{6}\text{C}^{*} \rightarrow {}^{12}_{6}\text{He} + X$. Determine the atomic and mass numbers for X and identify the type of radiation being emitted.

Worked example 3.2.4

RADIOACTIVE DECAY

Strontium-90 decays by radioactive emission to form yttrium-90. The equation is: ${}^{90}_{38}\text{Sr} \rightarrow {}^{90}_{39}\text{Y} + X$. Determine the atomic and mass numbers for *X* and identify the type of radiation being emitted.	
Thinking	**Working**
Balance the mass numbers.	The mass numbers of 90 are already balanced, so the mass number of *X* is zero.
Balance the atomic numbers.	$38 = 39 + Z$ $Z = 38 - 39 = -1$
X has an atomic number of –1 and a mass number of zero.	*X* is a beta-minus particle, ${}^{0}_{-1}\beta$.

Worked example: Try yourself 3.2.4

RADIOACTIVE DECAY

Polonium-218 decays by emitting an alpha particle and a gamma ray. The nuclear equation is:

${}^{218}_{84}\text{Po} \rightarrow X + {}^{4}_{2}\alpha + \gamma$.

Determine the atomic and mass numbers for *X*, then use the periodic table to identify the element.

EXTENSION

How radiation is detected

Our bodies cannot detect alpha, beta or gamma radiation. Therefore, a number of devices have been developed to detect and measure radiation.

A common detector is the Geiger counter. These are used:

- by geologists searching for radioactive minerals such as uranium
- to monitor radiation levels in mines
- to measure the level of radiation after a nuclear accident, such as the accident at Fukushima, Japan, in 2011
- to check the safety of nuclear reactors
- to monitor radiation levels in hospitals and factories.

A Geiger counter consists of a Geiger–Muller tube filled with argon gas, as shown in Figure 3.2.6.

A voltage of about 400 V is maintained between the positively charged central electrode and the negatively charged aluminium tube. When radiation enters the tube through the thin mica window, the argon gas becomes ionised and releases electrons. These electrons are attracted towards the central electrode and ionise more argon atoms along the way. For an instant, the gas between the electrodes becomes ionised enough to conduct a pulse of current between the electrodes. This pulse is registered as a count. The counter is often connected to a small loudspeaker so that the count is heard as a 'click' (Figure 3.2.7).

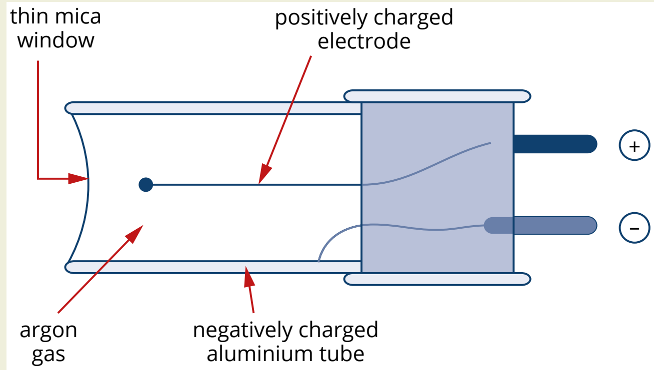


FIGURE 3.2.6 A schematic diagram of Geiger counter used for detecting ionising radiation.

FIGURE 3.2.7 A Russian scientist using a Geiger counter to measure radiation levels.

PHYSICS IN ACTION

How technetium is produced

Technetium-99m is the most widely used radioisotope in nuclear medicine. It is used for diagnosing and treating cancer. However, this radioisotope decays relatively quickly and so usually needs to be produced close to where it will be used. Technetium-99m is produced in small nuclear generators that are located in hospitals around the country (Figure 3.2.8). In this process, the radioisotope molybdenum-99, obtained from the Lucas Heights reactor, Sydney, is used as the parent nuclide. Molybdenum-99 decays by beta emission to form a relatively stable (or metastable) isotope of technetium, technetium-99m, as shown below:

$$^{99}_{42}\text{Mo} \rightarrow {}^{99\text{m}}_{43}\text{Tc} + {}^{0}_{-1}\beta + \nu$$

Technetium-99m is flushed from the generator using a saline solution. The radioisotope is then diluted and attached to an appropriate chemical compound before being administered to the patient as a tracer. Technetium-99m is purely a gamma emitter. This makes it very useful as a diagnostic tool for locating and treating cancer. Its decay equation is:

$$^{99\text{m}}_{43}\text{Tc} \rightarrow {}^{99}_{43}\text{Tc} + \gamma$$

FIGURE 3.2.8 Technetium generators are used in hospitals that require radioisotopes. The generator has a thick lead shield that absorbs the beta and most of gamma radiation.

3.2 Review

SUMMARY

- Radioactive isotopes may decay by emitting alpha, beta or gamma radiation from their nuclei.
- An alpha particle (α) consists of 2 protons and 2 neutrons and is emitted from the nuclei of some radioisotopes. It is identical to a helium nucleus and can be written as ${}^{4}_{2}He$.
- A beta-minus particle (β^- or ${}^{0}_{-1}\beta$) is an electron that has been emitted from the nucleus of a radioactive atom as a result of a neutron transmutating into a proton. An antineutrino ($\overline{v}$) is always emitted with a beta-minus particle.
- A beta-plus particle (β+) or positron is a positively charged electron (${}^{0}_{+1}\beta$) that has been emitted from the nucleus of a radioactive atom as a result of a proton transmutating into a neutron. A neutrino is always emitted with a beta-plus particle.
- A gamma ray (γ) is high-energy electromagnetic radiation that is emitted from the nuclei of radioactive atoms. It has no charge.
- In any nuclear reaction, both the atomic and mass numbers are conserved.

KEY QUESTIONS

1 Determine the nature of the unknown, *X*, for the following transmutation:
${}^{218}_{86}Rn \rightarrow {}^{214}_{84}Po + X$

2 In the following nuclear reaction, *Y* represents a beta particle. What type of beta particle is it?
${}^{214}_{82}Pb \rightarrow {}^{214}_{83}Bi + Y$

3 What type of decay occurs when a nucleus has too many protons?

4 What is a positron?

5 What types of radiation are alpha, beta and gamma radiation?

6 What are the mass numbers of the six stable nuclides of calcium (atomic number 20)? Use Figure 3.1.10 on page 51 to answer this question.

7 Where in an atom do alpha, beta and gamma radiation originate from?

8 For the unknown nuclides *X* and *Y* in each of these decay equations, determine the atomic number and mass number, and use the periodic table to identify the unknown elements.
a ${}^{235}_{92}U \rightarrow \alpha + X + \gamma$
b ${}^{228}_{88}Ra \rightarrow Y + \beta^- + \gamma$

9 Carbon-14 decays by beta-minus emission to form nitrogen-14. The equation for this is:
${}^{14}_{6}C \rightarrow {}^{14}_{7}N + {}^{0}_{-1}\beta + \overline{v} + \text{energy}$
a How many protons and neutrons does the nitrogen atom have?
b What particle on the left side of the equation has transformed into what particle(s) on the right side of the equation?

10 What is missing in each of the following decay equations?
a ${}^{45}_{20}Ca \rightarrow {}^{45}_{21}Sc + ?$
b ${}^{150}_{70}Yb \rightarrow {}^{146}_{68}Er + ?$

3.3 Properties of alpha, beta and gamma radiation

In the early experiments with radioactivity, emissions from a sample of radium were directed through a magnetic field as shown in Figure 3.3.1. The emissions followed three distinct paths, which suggested that there were three different forms of radiation being emitted. The emissions each had different charges, masses and speeds. These properties will be discussed in this section.

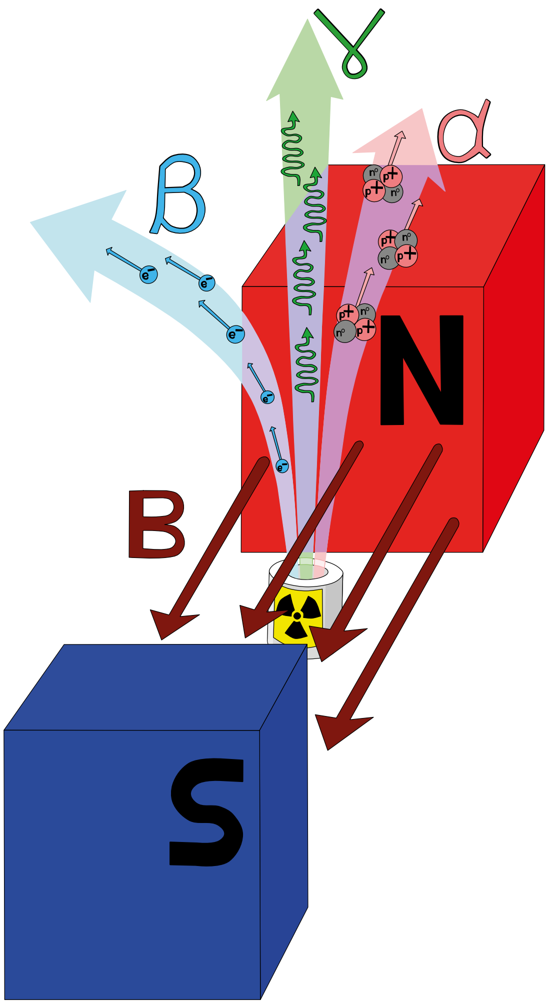


FIGURE 3.3.1 Applying a magnetic field shows that there are three different types of emissions from a radium source.

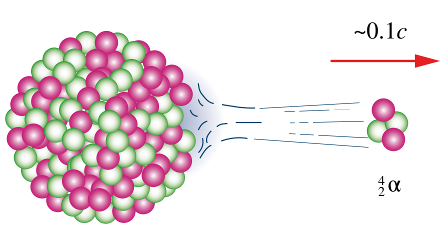


FIGURE 3.3.2 The speed and structure of an alpha particle being emitted from a nucleus.

SPEED AND CHARGE

Alpha (α) particles

Alpha particles consist of two protons and two neutrons. This means that they are relatively heavy and slow moving. Alpha particles are emitted from the nucleus at speeds of up to 20 000 km s^{-1} (2.0×10^{7} m s^{-1}), just less than 10% of the speed of light (Figure 3.3.2). Alpha particles have a double positive charge.

Beta (β) particles

Beta particles are fast-moving electrons (β^-), carrying a negative charge, or positrons (β^+), carrying a positive charge. Beta-minus particles are created when a neutron decays into a proton. Beta-plus particles are created when a proton decays into a neutron.

Beta particles are much lighter than alpha particles. As a result, they leave the nucleus with far higher speeds—up to 90% of the speed of light (c), as shown in Figure 3.3.3.

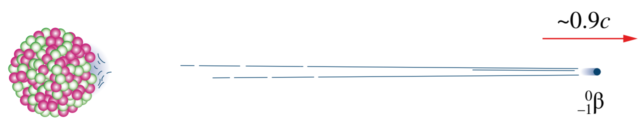


FIGURE 3.3.3 The speed and structure of a beta-minus particle being emitted from a nucleus.

Gamma (γ) rays

Gamma rays are electromagnetic radiation with very high frequency. Figure 3.3.4 shows where gamma rays lie along the electromagnetic spectrum. They have no rest mass and travel at the speed of light: $3.0 \times 10^8\,\mathrm{m\,s^{-1}}$ or $300\,000\,\mathrm{km\,s^{-1}}$ (Figure 3.3.5). Gamma rays have no electric charge.

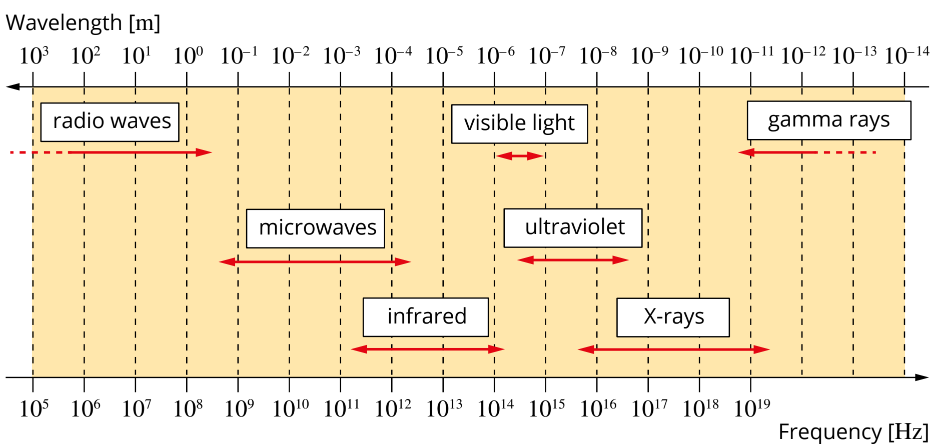


FIGURE 3.3.4 The electromagnetic spectrum contains many different types of radiation that differ in their wavelength and frequency. Gamma rays have very high frequencies and very short wavelengths, making them very energetic and highly penetrating.

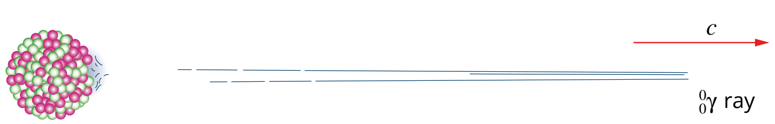


FIGURE 3.3.5 The speed and nature of gamma radiation.

PENETRATION ABILITY AND IONISATION

Alpha, beta, and gamma particles have widely different penetration ability in air and through solid objects due to the nature of the particles and their ionisation capability. Understanding their penetration ability allows us to design adequate shielding for radiation.

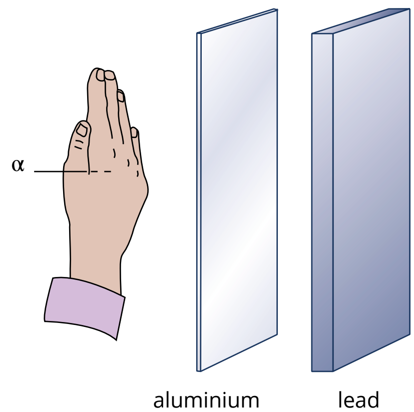


FIGURE 3.3.6 The penetrating ability of alpha radiation.

Alpha (α) particles

Alpha particles have relatively slow speed (approximately a tenth of the speed of light, or $0.1c$), heavy mass and a double charge. When an alpha particle travels through air, its slow speed and double positive charge cause it to interact with just about every atom that it encounters. The alpha particle dislodges electrons from many thousands of these atoms, turning them into ions—it ionises them. Each interaction slows the α particle down a little, and eventually it will be able to pick up some free electrons to become a helium atom. All this takes place within one or two centimetres in air. As a result, the air becomes quite ionised. The alpha particles are said to have a high **ionising ability**. Since the alpha particles do not get very far in the air, they have a poor **penetrating ability**.

Alpha particles only travel a few centimetres in air before losing their energy, and will be completely absorbed by thin card or a human hand (Figure 3.3.6).

An example of an isotope that emits alpha radiation (or undergoes alpha decay) is the isotope of americium $^{241}_{95}\text{Am}$. Americium can be found in ionisation smoke detectors (Figure 3.3.7). The alpha particles are easily absorbed within the detector.

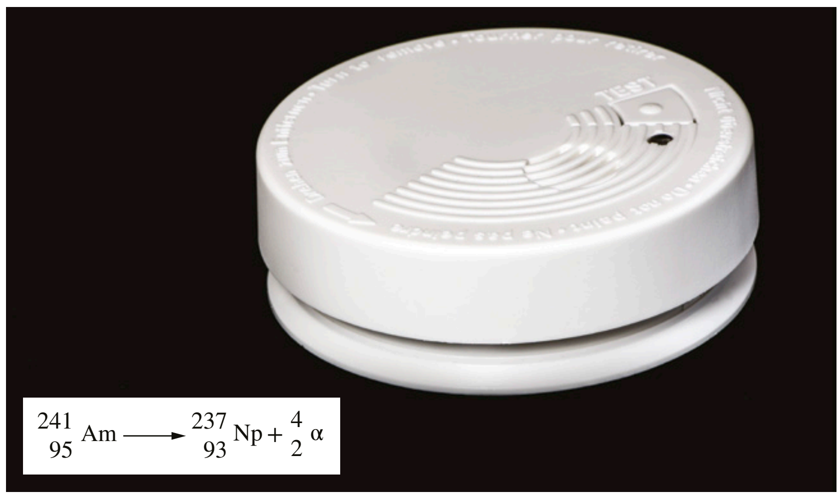


FIGURE 3.3.7 A domestic smoke detector contains a radioactive alpha emitter.

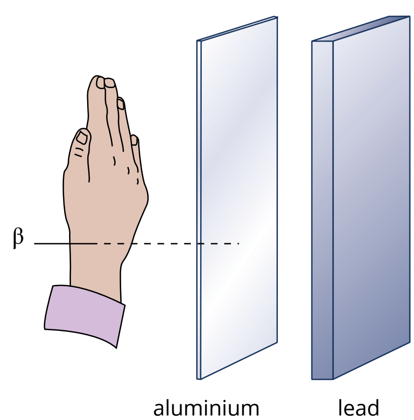


FIGURE 3.3.8 The penetrating ability of beta radiation.

Beta (β) particles

Beta particles are faster (approximately $0.9c$) than alpha particles and have a smaller charge (+1 or –1). Beta-minus particles have a negative charge and are repelled by the electron clouds of the atoms they interact with. This means that when a β^- particle travels through matter, it experiences a large number of glancing collisions. It loses less energy per collision than an α particle. As a result, β^- particles do not ionise atoms as readily and are more penetrating than α particles. Positrons, β^+ particles, are unstable and also have a weak penetrating power.

Beta particles will travel a few metres through air and through a human hand. Typically, a sheet of aluminium about 1 mm thick will stop them, as shown in Figure 3.3.8.

Gamma (γ) rays

Gamma rays have no charge and move at the speed of light. They are the most highly penetrating form of radiation. Gamma rays have a low probability of interacting with matter. This only occurs if they happen to collide directly with a nucleus or electron. The low density of an atom makes this relatively unlikely. Gamma rays pass through matter very easily—they have a very poor ionising ability but a high penetrating ability.

Gamma rays can travel an almost unlimited distance through air and even through a human hand, an aluminium sheet and a few centimetres of lead (Figure 3.3.9). A significantly thicker width of lead would be needed to fully stop the gamma rays. Even a metre of concrete would not completely absorb a beam of gamma rays.

PHYSICS IN ACTION

Monitoring the thickness of sheet metal

Beta-minus particles can be used to monitor the thickness of rolled sheets of metal and plastic during manufacture, as shown in Figure 3.3.11. A β^- particle source is placed under the newly rolled sheet and a detector is placed on the other side. If the sheet being made is too thick, fewer β^- particles will penetrate and the detector count will fall. This information is instantaneously fed back to the rollers and the pressure is increased until the correct reading is achieved and hence the right thickness of metal is attained.

Alpha particles or gamma rays would not be appropriate for this task.

Alpha particles have a very poor penetrating ability, so none would pass through the metal. Gamma rays usually have a high penetrating ability and so a thin metal sheet would not stop them. In addition, workers would need to be shielded from gamma radiation.

The penetrating properties of beta rays make them ideal for this job. The thickness of photographic film and plastic sheets is also monitored in this way.

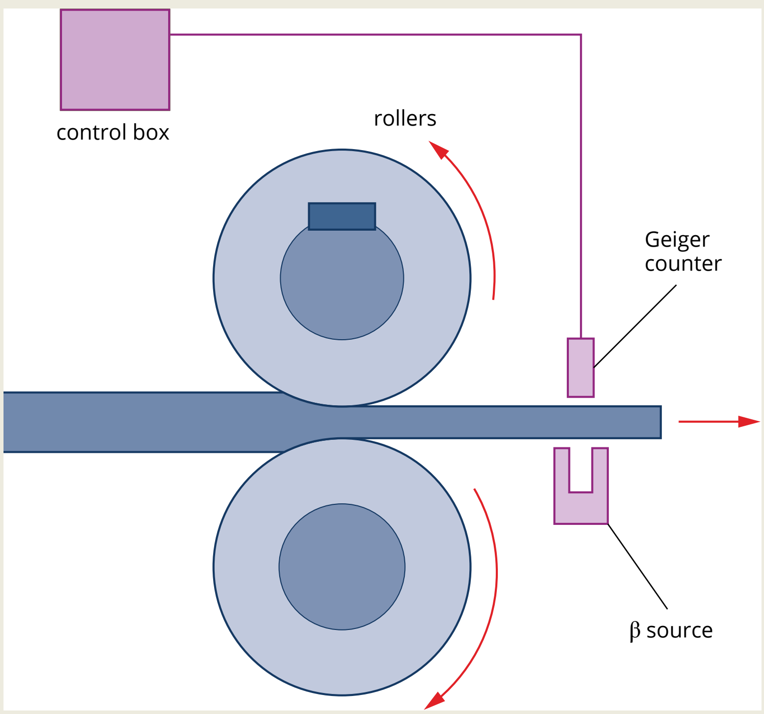


FIGURE 3.3.11 Beta emitters are used to monitor metal-sheet thickness.

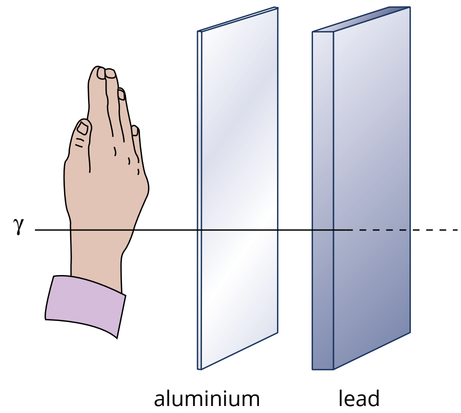


FIGURE 3.3.9 The penetrating ability of gamma radiation.

PHYSICSFILE

Gamma Knife Therapy

Gamma knife therapy is often used to target brain tumours without the need to make a surgical incision. Beams of gamma radiation (up to 201) placed at different angles around the head are programmed to intersect at the position of the tumour. Only the tumour being targeted receives a significant radiation dose, whilst the surrounding tissue is unharmed by the weaker individual beams.

FIGURE 3.3.10 A doctor prepares a patient for Gamma Knife Therapy.

EFFECT OF ELECTRIC AND MAGNETIC FIELDS

Electric fields are known to accelerate charged particles such as electrons and are often used in applications such as the Large Hadron Collider to accelerate charged particles to high speeds. Therefore, passing alpha, beta and gamma particles through an electric field will result in acceleration of the charged alpha and beta particles but will have no effect on the gamma rays as they have no charge.

A charged particle travelling through a magnetic field will have a force acting on it, at right angles to the direction of travel. This will cause the path of the charged particle to bend. The positively charged alpha particle will be deflected at right angles in one direction, and the negatively charged beta particle will be deflected in the opposite direction, also at right angles. (The details of this are covered in Year 12.) As shown in Figure 3.3.1 on page 62, the gamma particle will pass through the magnetic field without its path being changed.

ENERGY OF ALPHA, BETA AND GAMMA RADIATION

The energy of moving objects, such as cars and tennis balls, is measured in joules. However, alpha, beta and gamma radiation have such small amounts of energy that measuring it in joules is inappropriate. The energy of radioactive emissions is usually expressed in electronvolts (eV).

- An **electronvolt** is the energy that an electron would gain if it were accelerated by a potential difference of 1 volt.
- One electronvolt is an extremely small quantity of energy, equal to 1.6×10^{-19} J.

$1\ \text{eV} = 1.6 \times 10^{-19}\ \text{J}$

The range of energies for the different radiations are shown below.

- Alpha and beta particles are ejected from unstable nuclei with a wide range of energies.
- Alpha particles typically have energies of 5–10 million electronvolts or megaelectronvolts (5–10 MeV)
- Beta particles are usually ejected with energies up to a few million electronvolts or MeV.
- The speed of alpha and beta particles will increase as their energy increases.
- Gamma rays normally have less than 1 MeV of energy. For example, the gamma rays emitted by the radioactive isotope gold-198 have a maximum energy of 412 000 eV or 412 kiloelectronvolts (412 keV) or 6.6×10^{-6} J.
- The speed of a gamma ray does not increase as its energy increases. The gamma ray will still travel at the speed of light. The frequency of the gamma ray increases as the energy increases.

To give you an idea of how energetic these particles are, removing (ionising) an electron from the electron shell in hydrogen is 13.6 eV and any emission emitted from an electron relaxing back to its ground state is in the much lower energy visible spectra.

PROPERTIES OF ALPHA, BETA AND GAMMA RADIATION

The properties of alpha (α), beta (β) and gamma (γ) radiation are summarised in Table 3.3.1.

3.3.1 A comparison of the properties of alpha, beta and gamma radiation.

	alpha (α) particle	beta (β) particle	gamma (γ) ray
mass	heavy	light	none
speed	up to 20 000 km s^{-1} or about 10% of the speed of light	about 90% of the speed of light	the speed of light
charge	+2	–1 or +1	0
ionising ability	very high	low	very low
range in air	a few centimetres	1 or 2 m	many metres
penetration in matter	~10^{-2} mm	a few mm	high

3.3 Review

SUMMARY

- Alpha (α) particles are ejected from the nucleus at around 10% of the speed of light. They have a double positive charge and are relatively heavy. Alpha particles have high ionisation ability and poor penetrating power.
- Beta (β) particles are ejected from the nucleus at up to 90% of the speed of light. They are much lighter than alpha particles. β^- particles have a single negative charge. β^+ particles have a single positive charge. Beta radiation has moderate ionising and penetrating ability.
- Gamma (γ) rays are high-energy electromagnetic radiation and so travel at the speed of light and have no charge. They have high penetrating power and much lower ionisation ability.
- The alpha and beta particles can be accelerated by an electric field and have their paths bent by a magnetic field. The speed and direction of gamma rays are unaffected by electric and magnetic fields.

KEY QUESTIONS

1 Which type of radiation (alpha, beta or gamma):
  a can easily penetrate aluminium foil?
  b is ejected when a neutron decays into a proton?
  c travels relatively slowly at typically around 10% of the speed of light?
  d travels at speeds of up to 90% of the speed of light?
  e has no charge?

2 Which type of radiation (alpha, beta or gamma) is unaffected by a magnetic field?

3 Which type of radiation (alpha, beta or gamma) could penetrate human skin but not 1 mm of aluminium?

4 Which type of radiation (alpha, beta or gamma) would be used to treat brain cancer, where the radiation needs to penetrate the skull and reach the site of the tumour?

5 A radioactive source is emitting alpha, beta, and gamma radiation into the air. Which type/s of radiation would a Geiger counter held about 20 cm from the source most likely detect?

6 Where in the atom do the following types of radiation originate from?
  a alpha
  b beta
  c gamma

7 List the radiations alpha, beta and gamma in order of decreasing penetrating power.

8 Briefly explain why alpha particles have a very poor penetrating ability.

9 A radiographer inserts a radioactive wire into a breast cancer to destroy the cancerous cells close to the wire. Should this wire be an alpha, beta, or gamma emitter? Explain your reasoning.

10 Explain why beta particles, not alpha or gamma, are the best to use for monitoring the thickness of thin metal sheets.

3.4 Half-life and decay series

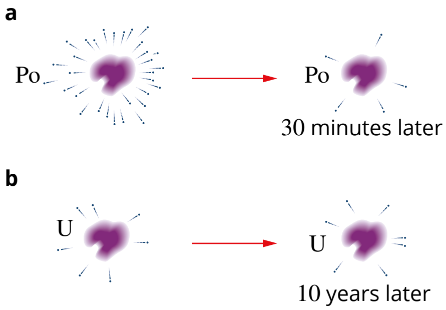


FIGURE 3.4.1 The activity of (a) polonium-218 and (b) uranium-235.

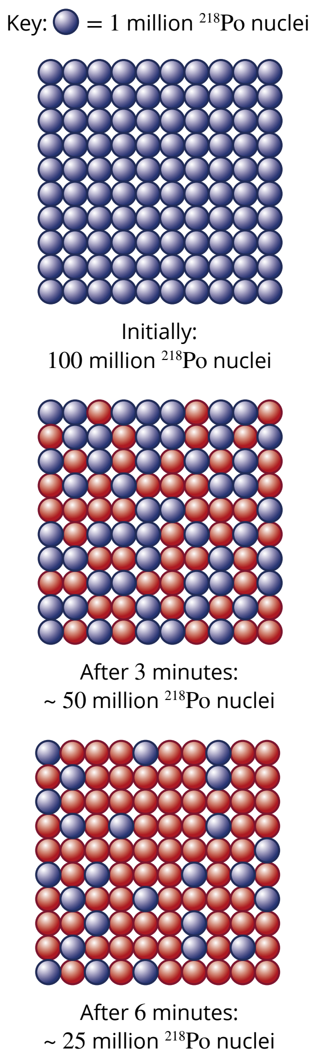


FIGURE 3.4.2 The decay of polonium-218 (blue dots) over two half-lives, showing how only one-quarter (25%) of the original radioisotope remains after two half-lives.

Different radioisotopes will emit radiation and will decay at very different rates. A Geiger counter held close to a small sample of polonium-218 will initially detect a very high level of radiation. Half an hour later, this count rate will have dropped to almost zero.

Compare this with a similar sample of uranium-235. A Geiger counter held close to the uranium will show a low count rate. However, as time passes, the count rate does not seem to change. If you came back decades later, the count level would still be the same. Figure 3.4.1 compares the activity of both of these radioisotopes.

The half-life of a radioisotope describes how long it takes for half of the atoms in a given mass to decay. The count rate is the activity of the sample. These ideas will be studied in this section.

HALF-LIFE

All radioisotopes are unstable but some are more unstable than others. In the previous example, polonium-218 is more unstable than uranium-235. One way of determining this instability is by measuring the **half-life** ($t_{1/2}$) of the radioisotope.

> The half-life ($t_{1/2}$) of a radioisotope is the time that it takes for half of the nuclei of the sample radioisotope to decay.

The half-life of polonium-218 is 3 minutes. Consider a sample of polonium that initially contains 100 million undecayed polonium-218 nuclei, as shown in Figure 3.4.2. Over the first 3 minutes about half of these will have decayed, leaving around 50 million polonium-218 nuclei. Over the next 3 minutes, half of these 50 million polonium-218 nuclei will decay, leaving approximately 25 million of the original radioactive nuclei. The process continues as time passes.

> The number of nuclei remaining after a particular number of half-lives can be found mathematically using:
>
> $$N = N_0 \, \frac{1}{2}^{\,n}$$
>
> where N is the number of radioactive nuclei remaining
> N_0 is the initial number of radioactive nuclei
> n is the number of half-lives elapsed.
>
> The number of half-lives in a period of time can be found using:
>
> $$n = \frac{T}{t_{1/2}}$$
>
> where n is the number of half-lives elapsed
> T is the period of time that the radioactive nuclei has decayed
> $t_{1/2}$ is the half-life of the radioactive nuclei.

As time passes, a smaller and smaller proportion of the original radioisotope remains in the sample, until eventually the amount of decay is negligible. The graph in Figure 3.4.3, known as a decay curve, shows this.

Even a very small radioactive sample will contain billions of atoms. It is important to know that although the behaviour of such a large sample of nuclei can be mathematically predicted, the behaviour of one particular nucleus cannot. It has a 50% chance of decaying in each half-life. Also, the half-life of a radioisotope is constant and cannot be changed by chemical reactions, heat and so on. Half-life is solely determined by the instability of the nuclei of the radioisotope.

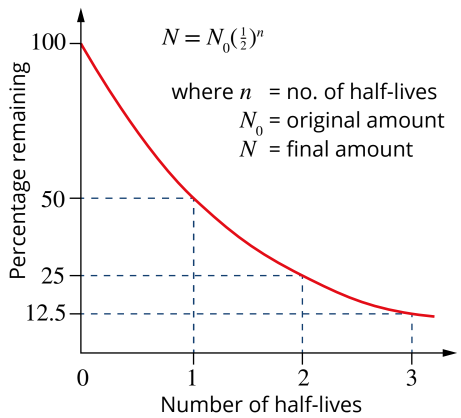


FIGURE 3.4.3 A decay curve for a radioisotope.

Worked example 3.4.1

HALF-LIFE

A sample of the radioisotope thorium-234 contains 8.0×10^{12} nuclei. The half-life of thorium-234 is 24 days. How many thorium-234 nuclei will remain in the sample after 120 days?	
Thinking	**Working**
Calculate how many half-lives 120 days corresponds to.	$n = \frac{120}{24}$ $= 5$ half-lives
Substitute $N_0 = 8.0 \times 10^{12}$ and $n = 5$ into the equation. Calculate the number of nuclei remaining.	$N = N_0 \left(\frac{1}{2}\right)^n$ $= 8.0 \times 10^{12} \times \left(\frac{1}{2}\right)^5$ $= 2.5 \times 10^{11}$ nuclei

Worked example: Try yourself 3.4.1

HALF-LIFE

A sample of the radioisotope sodium-24 contains 4.0×10^{10} nuclei. The half-life of sodium-24 is 15 hours. How many sodium-24 atoms will remain in the sample after 150 hours?

ACTIVITY

A Geiger counter can be used to record the number of radioactive decays occurring in a sample each second. It measures the **activity** of the sample.

Activity is measured in becquerels, Bq.

1 Bq = 1 disintegration per second.

$$A = \frac{N}{t}$$

where A is the activity

N is the number of disintegrations

t is the time

Over time, the activity of a sample of a radioisotope will decrease. More and more of the radioactive nuclei will have decayed and at some point it will no longer emit radiation.

If a sample of polonium-218 ($t_{1/2}$ = 3 minutes) has an initial activity of 2000 Bq, then after one half-life its activity will be 1000 Bq. After a further 3 minutes, the activity of the sample will have reduced to 500 Bq, and after a further 3 minutes it will be 250 Bq and so on.

Uranium-235 has a half-life of 700 000 years. Its activity will remain virtually constant for decades and will certainly not change over 3 minutes.

The equation used to find the number of nuclei after a number of half-lives can also be used to calculate the final activity of a radioactive sample after a number of half-lives.

The activity of the nuclei remaining after a number of half-lives can be found mathematically using:

$$A = A_0 \left(\frac{1}{2}\right)^n$$

where A is the activity of the radioactive nuclei remaining

A_0 is the initial activity of the radioactive nuclei

n is the number of half-lives elapsed.

Worked example 3.4.2

HALF-LIFE AND ACTIVITY

Technetium-99m is a metastable compound with a half-life of 6 hours that is often used for radioactive imaging of bones, the brain and the lungs. If a sample of technetium-99m has an initial activity of 4000 Bq, what will its activity be after 2 days?

Thinking	Working
Calculate how many half-lives 2 days corresponds to. Convert the number of days to hours first.	2 days = 2 × 24 = 48 hours $n = \frac{48}{6} = 8$ = 8 half-lives
Substitute the initial activity A_0 = 4000 and the number of half-lives, n = 8, into the equation. Calculate the final activity	$A = A_0 \left(\frac{1}{2}\right)^n$ $= 4000 \times \left(\frac{1}{2}\right)^8$ = 16 Bq

You will notice that due to the short half-life the activity has significantly reduced after 8 half-lives or 2 days, making it ideal for medical purposes.

Worked example: Try yourself 3.4.2

HALF-LIFE AND ACTIVITY

A sample of strontium-90 has an initial activity of 4000 Bq. Calculate its activity after 6 months using Table 3.4.1.

COMMON RADIOISOTOPES AND THEIR APPLICATIONS

The half-lives of some common radioisotopes are shown in Table 3.4.1. The half-life of a radioisotope will determine what it is used for. For example, the most commonly used medical tracer, technecium-99, has a short half-life of just 6 hours. The short half-life means that radioactivity does not remain in the body any longer than necessary. On the other hand, the radioisotope used in a smoke detector, americium-241, is chosen because of its long half-life, 461 years. The smoke detector can continue to function for a very long time, as long as the battery is replaced each year.

TABLE 3.4.1 Some common radioisotopes, their half-lives and applications.

Isotope	Emission	Half-life	Application
Natural			
polonium-214	α	0.000 16 s	nothing at this time
strontium-90	β	28.8 years	cancer therapy
radium-226	α	1630 years	once used in luminous paints
carbon-14	β	5730 years	carbon dating of fossils
uranium-235	α	700 000 years	nuclear fuel, rock dating
uranium-238	α	4.5 billion years	nuclear fuel, rock dating
thorium-232	α	14 billion years	fossil dating, nuclear fuel
Artificial			
technetium-99m	β	6 h	medical tracer
sodium-24	β	15 h	medical tracer
iodine-131	β	8 days	medical tracer
phosphorus-32	β	14.3 days	medical tracer
cobalt-60	β	5.3 years	radiation therapy
americium-241	α	460 years	smoke detectors
plutonium-239	α	24 000 years	nuclear fuel, rock dating

DECAY SERIES

Generally, when a radionuclide decays, its daughter nucleus is not completely stable and is itself radioactive. This daughter nucleus will then undergo further decay. Eventually a stable isotope is reached and the sequence ends. This is known as a **decay series**. An example of a decay series is shown in Figure 3.4.4. This particular series shows the decay of uranium-238 (shown at the top of the chart). The uranium-238 has a long half-life of 4.5×10^9 years but eventually decays into thorium-234 by alpha emission. Thorium-234 has a short half-life of only 24 days, decaying into protactinium 234 by beta decay, and so on until the final product is lead-206 (shown at the bottom of the chart).

The Earth is 4.5 billion years old—old enough to have only four naturally occurring decay series that remain active. These are:

- the uranium series in which uranium-238 eventually becomes lead-206
- the actinium series in which actinum-235 eventually becomes lead-207
- the thorium series in which thorium-232 eventually becomes lead-208
- the neptunium series in which neptunium-237 eventually becomes bismuth-209. (Since neptunium-237 has a relatively short half-life, it is no longer present in the crust of the Earth, but the rest of its decay series is still continuing.)

Geologists analyse the proportions of the radioactive elements in a sample of rock to gain a reasonable estimate of the rock's age. This technique is known as rock dating.

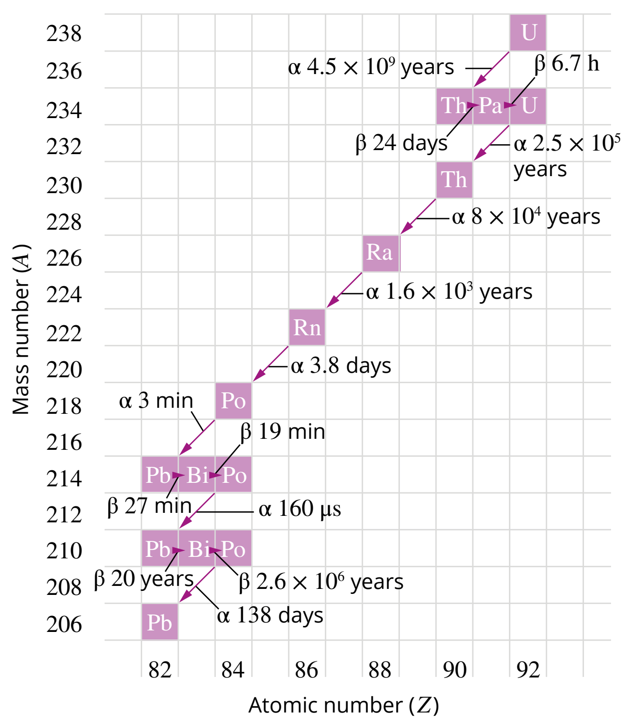


FIGURE 3.4.4 The uranium decay series. The half-life and emissions are indicated on each of the decays as radioactive uranium-238 is gradually transformed into stable lead-206. Mining companies find significant quantities of lead at uranium mines.

EXTENSION

Radiocarbon dating

Radiocarbon dating (or carbon dating) is a technique used by archaeologists to determine the ages of fossils and ancient objects that were made from plant matter. This method involves measuring and comparing the proportion of two isotopes of carbon, carbon-12 and carbon-14, in the specimen.

Carbon-12 is a stable isotope but carbon-14 is radioactive. Carbon-14 only exists in trace amounts in nature. Carbon-12 atoms are about 1 000 000 000 000 (10^{12}) times more common than carbon-14 atoms.

Carbon-14 has a half-life of 5730 years and decays by β^- emission to nitrogen-14. Its decay equation is:

$${}^{14}_{6}\mathrm{C} \rightarrow {}^{14}_{7}\mathrm{N} + \beta^- + \bar{\nu}$$

Both carbon-12 and carbon-14 can combine with other atoms in the environment. For example, they both combine with oxygen to form carbon dioxide. Plants and animals take in carbon-based molecules from the air and food. This means that all living organisms contain the same percentage of carbon-14. In the environment, the production of carbon-14 is matched by its decay and so the proportion of carbon-14 atoms to carbon-12 remains constant.

After an organism dies, the amount of carbon-14 it contains will decrease as these atoms decay to form nitrogen-14 and are not replaced from the environment. The number of atoms of carbon-12 does not change as carbon-12 is a stable atom. So, over time, the proportion of carbon-14 to carbon-12 atoms decreases.

The proportion of carbon-14 to carbon-12 in a dead organism can be compared with that found in living organisms and the approximate age of the specimen can be determined from the half-life of carbon-14.

Consider this example. The count rate from a 1 gram sample of carbon that has been extracted from an ancient wooden spear is 10 Bq. A 1 gram sample of carbon from a living piece of wood has a count rate of 40 Bq. We can assume that this was also the initial count rate of the spear. For its count rate to have reduced from 40 to 10 Bq, the spear must be (40 → 20 → 10) two half-lives of carbon-14 old, therefore with carbon-14 having a half-life of 5730 years, the spear is about 11 500 years old.

In 1988, scientists used carbon-dating techniques to show that the Shroud of Turin was probably a medieval forgery. It had been claimed that the Shroud of Turin was the piece of cloth that was the burial shroud of Jesus Christ (Figure 3.4.5). Carbon-dating tests on small samples of the cloth established that there was a high probability that it was made between 1260 and 1390, not around the time of Christ's death.

FIGURE 3.4.5 The Shroud of Turin.

Radiocarbon dating is an important aid to anthropologists who are interested in finding out about the migration patterns of early people—including the Australian Aborigines. This technique is very powerful since it can be applied to the remains of ancient campfires. It is accurate and reliable for samples up to about 60 000 years old. Carbon dating cannot be used to date dinosaur bones as they are more than 60 million years old, but it can be used to determine the age of more recently extinct mammoth fossils, like that shown in Figure 3.4.6.

FIGURE 3.4.6 A fossilised mammoth analysed by carbon dating.

3.4 Review

SUMMARY

- The rate of decay of a radioisotope is measured by its half-life, $t_{1/2}$. This is the time that it takes for half of the radioisotope to decay.
- The activity of a sample indicates the number of emissions per second. Activity is measured in becquerels (Bq), where 1 Bq = 1 emission per second.
- The number of atoms of a radioisotope will decrease over time. Over one half-life, the number of atoms of a radioisotope will halve.
- The half-life equation can be used to calculate the number (N) or activity (A) of a radioisotope remaining after a number of half-lives (n) has passed:

$$N = N_0 \left(\frac{1}{2}\right)^n, A = A_0 \left(\frac{1}{2}\right)^n$$

- When a radionuclide decays, its daughter nucleus is usually itself radioactive. This daughter will then decay to a grand-daughter nucleus, which may also be radioactive, and so on. This is called a decay series.

KEY QUESTIONS

1 What is meant by the activity of a radioisotope?

2 Technetium-99m has a half-life of 6.0 hours. A sample of the radioisotope originally contains 8.0×10^{10} atoms. How many technetium-99m nuclei remain after 6.0 hours?

3 Iodine-131 has a half-life of 8 days. A sample of the radioisotope initially contains 2.4×10^{12} iodine-131 nuclei. How many iodine-131 nuclei remain after 24 days?

4 Radioactive materials are considered to be relatively safe when their activity has fallen below 0.1% of the initial value.

a How many half-lives does this take?

b Plutonium-239 is a by-product of nuclear reactors. Its half-life is 24000 years. How long does the plutonium-239 have to be stored as nuclear waste before it is considered safe to handle?

5 If a particular atom in a radioactive sample has not decayed during the previous half-life, what is the percentage chance that it will decay in the next half-life?

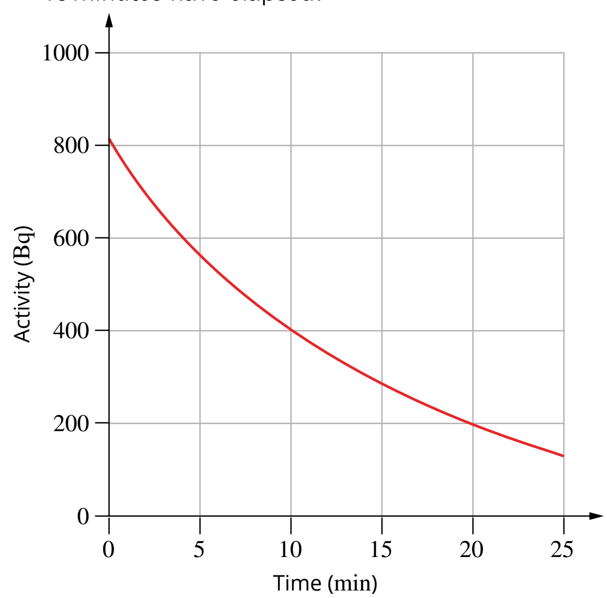
6 A hospital in Alice Springs needs 12 µg of the radioisotope technetium-99m. The specimen has to be ordered from Sydney. The half-life of technetium-99m is 6 hours and the delivery takes 24 hours. How much must be produced in Sydney to satisfy the Alice Springs order?

7 The activity of a radioisotope changes from 6000 Bq to 375 Bq over a period of 60 minutes. Calculate the half-life of this radioisotope.

8 A Geiger counter is used to measure the radioactive emissions from a certain radioisotope. The activity of the sample is shown in the graph.

a What is the half-life of the radioisotope according to the graph?

b What would the activity of the sample be after 40 minutes have elapsed?



9 According to Figure 3.4.4 on page 72, what type of decay does lead-210 undergo and what is its half-life?

10 In the uranium decay series shown in Figure 3.4.4, $^{234}_{92}U$ decays to eventually produce stable $^{206}_{82}Pb$. How many alpha and beta-minus decays have occurred?

3.5 Radiation doses and effects on humans

We are exposed to low levels of ionising radiation, known as background radiation, throughout our lives. However, exposure to high levels of ionising radiation from alpha, beta and gamma sources is harmful to humans and other living organisms. People who work with radiation in fields such as medicine, mining, nuclear power plants and industry must be able to closely monitor the radiation dose to which they are exposed. Similarly, radiologists who administer courses of radiation treatment to cancer patients need to be able to measure the radiation dose that they are applying. Figure 3.5.1 shows a dosimeter used by doctors, radiologists and technicians who work with gamma radiation to monitor their exposure levels. In this section, you will learn about radiation doses, how they are measured, the effects of high exposure and medical applications of radiation.

FIGURE 3.5.1 A dosimeter used to monitor gamma radiation exposure.

IONISING RADIATION

The electromagnetic spectrum consists of a variety of types of electromagnetic radiation. From low energy to high energy, these are: radio waves, microwaves, infrared radiation, visible light, ultraviolet radiation, X-rays and gamma rays, and are shown in Figure 3.5.2. **Non-ionising** radiation includes radio waves, microwaves, infrared radiation, visible light and UV-A radiation. Every day, people are exposed to significant amounts of such radiation without serious consequences.

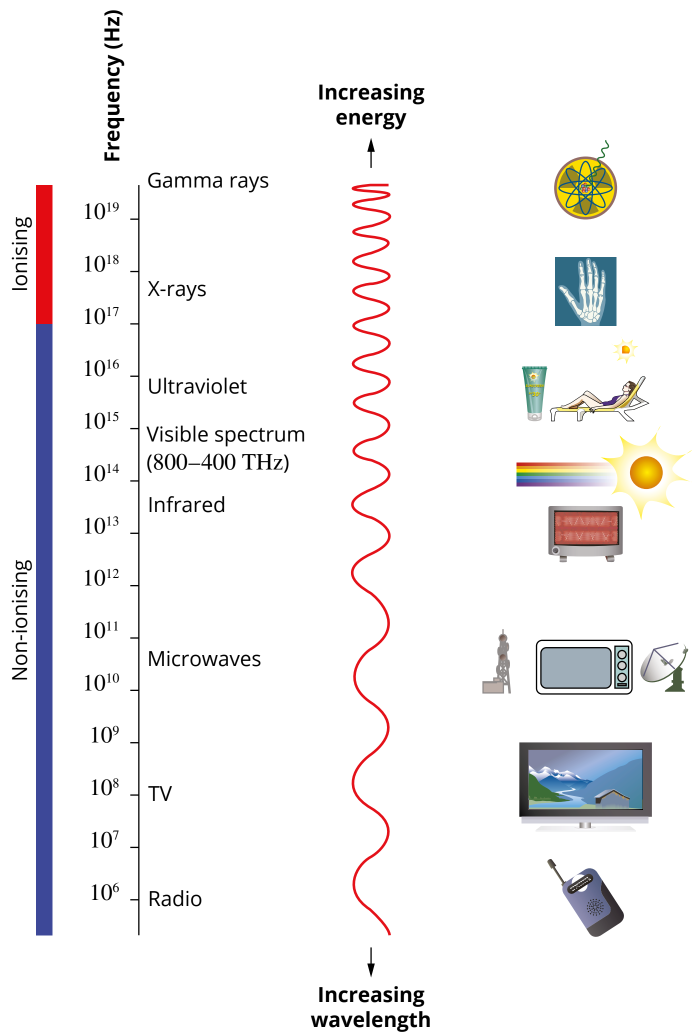


FIGURE 3.5.2 The electromagnetic spectrum consists of ionising and non-ionising radiation.

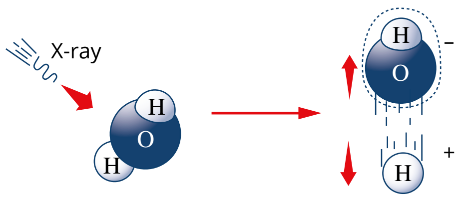


FIGURE 3.5.3 Water molecule being ionised.

Some forms of electromagnetic radiation are harmful to living things. These ionising radiations are found at the high-energy end of the electromagnetic spectrum, starting at UV-C, which is blocked by the ozone layer, and include gamma and X-rays. **Ionising radiation** is electromagnetic radiation with a frequency above 2×10^{16} Hz. When ionising radiation interacts with the tissue in an organism, it creates ions, which can damage tissue or lead to the development of cancerous tumours. Figure 3.5.3 shows an X-ray heading towards a water molecule. The ionising radiation has enough energy to break the bonds within the water molecule and create a pair of ions.

The Earth itself is radioactive and we are all exposed to a low level of ionising radiation from many different sources, as shown in Figure 3.5.4. Almost 90% of our annual exposure is from the surrounding environment. This radiation is called **background radiation** and this background level of exposure is not a significant problem to our health.

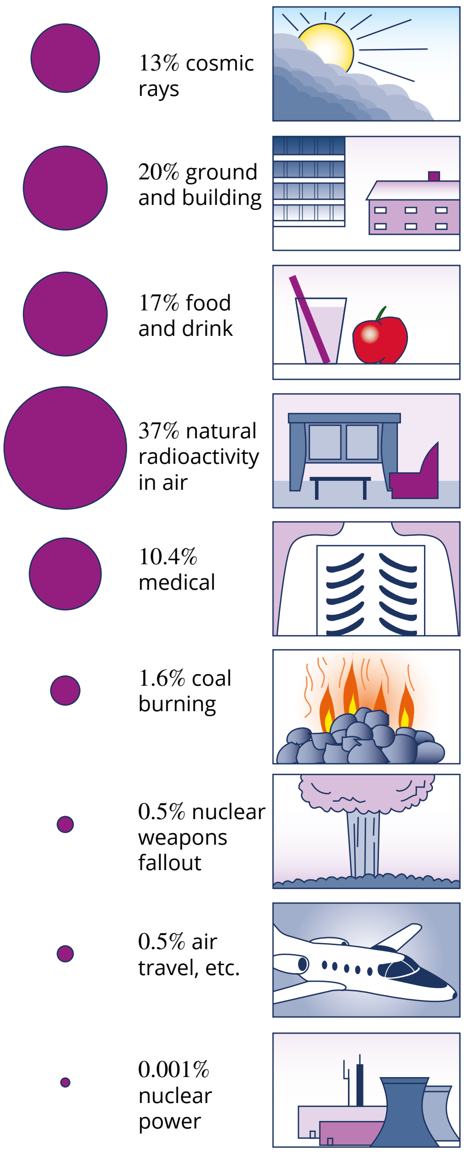


FIGURE 3.5.4 Everyone is exposed to radiation from the environment.

MEASURING RADIATION EXPOSURE

Exposure to high-energy radiation is harmful to living tissue. The energy of the radiation acts to break apart molecules and ionise atoms in the body's cells. It may lead to long-term problems such as cancer and deformities in future generations. Extremely high levels of radiation exposure can cause death, and this can happen within just a few hours. It is therefore important to be able to measure the amount of exposure a person has had.

Absorbed dose

The severity of radiation exposure depends on the amount of radiation energy that has been absorbed (E) and the mass of tissue involved (m).

The radiation energy absorbed per kilogram of tissue is called the **absorbed dose** (AD).

$$\text{absorbed dose} = \frac{\text{energy absorbed by the tissue}}{\text{mass of tissue}}$$

$$AD = \frac{E}{m}$$

Absorbed dose is measured in joules per kilogram (J kg^{-1}) or grays (Gy), i.e. 1 Gy = 1 J kg^{-1}.

Worked example 3.5.1

ABSORBED DOSE

A cancer tumour of mass 150 g is exposed to 0.30 J of radiation energy. Calculate the absorbed dose (AD) in grays. Assume that all of the radiation is absorbed by the tumour.

Thinking	Working
Convert the mass into kg.	150 g = 0.150 kg
Use $AD = \frac{E}{m}$ to calculate the absorbed dose.	$AD = \frac{0.30}{0.150}$ $= 2$ Gy

Worked example: Try yourself 3.5.1

ABSORBED DOSE

A cancer tumour is exposed to 0.50 J of radiation energy. The absorbed dose is 3 Gy. Calculate the mass of the tumour. Assume that all of the radiation is absorbed by the tumour.

Dose equivalent

Absorbed dose is not widely used when measuring radiation dose. It does not take into account the type of radiation involved. **Dose equivalent** (DE) does and it is the most common way in which radiation doses are measured.

Alpha particles are the most ionising form of radiation. This is because of their low speed, high charge and large mass (discussed in Section 3.3). Alpha particles interact with, and ionise, almost every atom that lies in their path. This means that an absorbed dose of alpha radiation is about 20 times more damaging to human tissue than an equal absorbed dose of beta or gamma radiation.

The weighting of the biological impact of radiation is called the **quality factor**, QF, or weighting factor. A list of quality factors is shown in Table 3.5.1

TABLE 3.5.1 The quality factors for various types of radiation.

Radiation	Quality factor
α-particles	20
neutrons* (10 keV)	10
β-particles	1
γ-rays	1
X-rays	1

*Radiation from neutrons is only found around nuclear reactors and neutron bomb explosions.

Gamma rays and X-rays have relatively low ionising powers. They have no charge and move at the speed of light so they fly straight past most atoms and interact only occasionally as they pass through. Gamma rays and X-rays would cause only slight damage to living cells. Beta-minus particles are considered to be as damaging as gamma rays and X-rays. This is reflected in their low quality factor of 1.

Neutrons are a more damaging form of radiation and have a weighting factor of 10. Neutrons are present in fission reactions which take place in nuclear reactors and bombs (fission will be covered in Chapter 4). The high quality factor explains why accidents at nuclear reactors are so dangerous to workers who are present.

The dose equivalent takes into account the absorbed dose and the type of radiation. This gives a more accurate picture of the actual effect of the radiation on tissue.

dose equivalent = absorbed dose × quality factor

This can be abbreviated to:

$DE = AD \times QF$

Dose equivalent is measured in Sieverts (Sv).

An absorbed dose of just 0.05 Gy of alpha radiation is equally as damaging to a person as an absorbed dose of 1.0 Gy of beta radiation. While less energy is carried by the α-particles than the β-particles, each α-particle does far more damage. For the 0.05 Gy alpha particle the $DE = 0.05 \times 20 = 1$ Sv and for the 1.0 Gy beta particles the $DE = 1.0 \times 1 = 1$ Sv. In each case, the dose equivalent is 1 Sv, and 1 Sv of any radiation causes the same amount of damage.

Worked example 3.5.2

ABSORBED DOSE

Calculate the dose equivalent (in μSv) from various types of radiation if the absorbed dose is 0.50 μGy.	
Thinking	**Working**
The quality factor for each type of radiation can be found in Table 3.5.1. 1 μGy = 1×10^{-6} Gy	QF (alpha particles) = 20 QF (beta particles) = 1 QF (gamma rays) = 1

a Calculate the dose equivalent (in μSv) from a radiation source if the absorbed dose is 0.50 μGy and the source is emitting alpha particles.	
The dose equivalent, DE = AD × QF.	$DE\ (\alpha) = 0.50 \times 10^{-6} \times 20$ $= 1 \times 10^{-5}$ Sv = 10 μSv

b Calculate the dose equivalent (in μSv) from a radiation source if the absorbed dose is 0.50 μGy and the source is emitting beta particles.	
Thinking	**Working**
The dose equivalent, DE = AD × QF.	$DE(\beta) = 0.50 \times 10^{-6} \times 1$ $= 0.50 \times 10^{-6}$ = 0.50 μSv

c Calculate the dose equivalent (in μSv) from a radiation source if the absorbed dose is 0.50 μGy and the source is emitting gamma rays.	
Thinking	**Working**
The dose equivalent, DE = AD × QF.	$DE(\gamma) = 0.50 \times 10^{-6} \times 1$ $= 0.50 \times 10^{-6}$ = 0.50 μSv

Worked example: Try yourself 3.5.2

DOSE EQUIVALENT

Calculate the dose equivalent (in mSv) from various radiation sources if the absorbed dose is 1.25 mGy.

a Calculate the dose equivalent (in mSv) from a radiation source if the absorbed dose is 1.25 mGy and the source is emitting alpha particles.

b Calculate the dose equivalent (in mSv) from a radiation source if the absorbed dose is 1.25 mGy and the source is emitting beta particles.

c Calculate the dose equivalent (in mSv) from a radiation source if the absorbed dose is 1.25 mGy and the source is emitting gamma rays.

Worked example: 3.5.3

TREATING TUMOURS

A 10 g cancer tumour absorbs 2.5×10^{-3} J of energy from an applied radiation source. Calculate the dose equivalent if the source is an alpha emitter using information from Table 3.5.1.	
Thinking	**Working**
Convert mass from grams to kg.	$m = \frac{10}{1000}$ $= 0.010\,\text{kg}$
Calculate the absorbed dose (AD) using the energy and the mass.	$\text{AD} = \frac{E}{m}$ $= \frac{2.5 \times 10^{-3}}{0.010}$ $= 0.25\,\text{Gy}$
Calculate the dose equivalent (DE) using the quality factor for alpha particles of 20.	$\text{DE} = \text{AD} \times \text{QF}$ $= 0.25 \times 20$ $= 5\,\text{Sv}$

Worked example: Try yourself 3.5.3

TREATING TUMOURS

A 25 g cancer tumour absorbs 5.0×10^{-3} J of energy from an applied radiation source. Calculate the dose equivalent if the source is an alpha emitter using information from Table 3.5.1.

People who work in occupations that involve ongoing exposure to levels of ionising radiation usually pin a small radiation-monitoring device to their clothing. This is usually a thermoluminescent dosimeter (TLD), as pictured in Figure 3.5.5. Thermoluminescent dosimeters contain a disc of lithium fluoride encased in plastic. Lithium fluoride can detect beta and gamma radiation, as well as X-rays and neutrons. TLDs are a cheap and reliable method for measuring radiation doses.

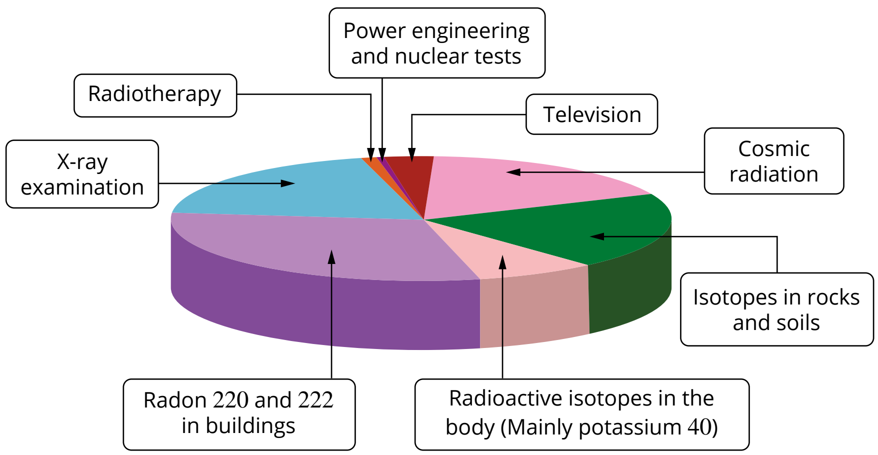
Dosimeters like these are used by personnel in nuclear power plants, radiotherapy departments at hospitals, airport security gates and uranium mines.

FIGURE 3.5.5 Thermoluminescent dosimeters are used by doctors, radiologists and scientists who work with radiation to monitor their exposure levels.

BACKGROUND RADIATION

It is important to appreciate that 1 Sv is a massive dose of radiation. It would not be fatal, but it would certainly lead to a severe case of radiation sickness. However, we are exposed to background radiation continually and it is not harmful to us. In Australia, the average annual background radiation dose is about 1.5 mSv (millisievert) or 1500 μSv. The average in the world is 2.4 mSv, and it is 3 mSv in North America. Figure 3.5.6 shows the origin of background radiation. In India there is a village situated above natural uranium and thorium deposits. The average equivalent dose for its inhabitants is 15 mSv per year, three times the recommended equivalent dose.

Share of the different sources of radiation
The average equivalent of the dose absorbed by a person in a year is equal to a few millisieverts.



Natural sources	*DE* [mSv]	Artificial sources	*DE* [mSv]
Cosmic radiation	0.3 – 0.5	X-ray examination	0.5
Isotopes in rocks and soils	0.4 – 0.5	Radiotherapy	0.01
Radioactive isotopes in the body (Mainly potassium 40)	0.2	Power engineering and nuclear tests	0.01
Radon 220 and 222 in buildings	0.6 – 0.8	Television	0.08
Total	1.5 – 2.0	Total	0.6

FIGURE 3.5.6 Pie chart showing common sources of background radiation.

Table 3.5.2 can be used to estimate your own dose over the past year.

TABLE 3.5.2 Annual radiation doses in Australia for different radiation sources.

Radiation source	Average annual dose (μSv)	Local variations
cosmic radiation	300	Plus 200 μSv for each round-the-world flight Plus 20 μSv for each 10° of latitude Plus 150 μSv if you live 1000 m above sea level
rocks, air and water	1350	Plus 1350 μSv if you live underground Plus 1350 μSv if your house is made of granite Minus 140 μSv if you live in a weatherboard house
radioactive foods and drinks	350	Plus 1000 μSv if you have eaten food affected by the Fukushima fallout
manufactured radiation	60	Plus 60 μSv if you live near a coal-burning power station Plus 30 μSv from nuclear testing in the Pacific

EFFECT OF RADIATION EXPOSURE ON HUMANS

High levels of radiation exposure can have an adverse effect on human tissue. Radiation exposure can be chronic (occurs over a long period of time), as would be expected by a worked in a nuclear reactor plant, or acute (occurs all at once), such as in in a nuclear reactor accident or explosion. Immediate effects of radiation exposure include a drop in the white blood cell count, nausea, fatigue, hair-loss and skin reddening. Radiation has the potential to cause DNA damage and cause cancerous cells. In addition, different radioactive elements can target different parts of the body—for example iodine affects the thyroid and strontium-90 affects the bone and bone marrow.

Table 3.5.3 shows some typical levels of radiation dosages and their effects, as put out by the World Nuclear Association. The effects discussed are assuming the whole body has been exposed. Dosages of radiation for cancer treatment can be significantly higher but these are specifically targeted at the cancerous cells.

TABLE 3.5.3 Effect of radiation dosage on the on the human body.

Radiation doses in millisieverts (mSv)	Effects
1.5 mSv	Typical background exposure in Australia
9 mSv/year	Exposure by airline crew on the New York–Tokyo route
100 mSv/year	Highest annual safe level; above this the probability of cancer is assumed to increase with the dose 4 months on the International Space Station, 350 km above the Earth
350 mSV/year	Criterion for relocating people after the Chernobyl accident
700 mSv/year	Suggested threshold for maintaining evacuation after a nuclear incident
1000 mSv short term, whole body	Threshold for causing radiation sickness and nausea, but not death
5000 mSv whole body	Fatal for 50% of those exposed
10000 mSv whole body	Acute radiation poisoning, death within a few weeks

MEDICAL APPLICATIONS OF RADIATION

Rapidly dividing cancerous cells are more susceptible to radiation than healthy tissue, so radiation is often used in cancer treatment. In addition, sources such as cobalt-60 can be used to specifically target only the cancerous tissue, minimising the damage to healthy tissue as shown in Figure 3.5.8.

FIGURE 3.5.8 A cancer patient about to receive radiation therapy.

PHYSICSFILE

Iodine-131 and the thyroid

On 11th March 2011, a catastrophic earthquake and tsunami hit Japan, killing tens of thousands of people and severely damaging the nuclear power station at Fukushima. Radioactive materials, including caesium-137 and iodine-131, escaped. These have half-lives of 30 years and 8 days respectively.

Our bodies need iodine for the healthy functioning of the thyroid gland, which maintains a proper metabolism. Foods rich in iodine include seafood, vegetables and salt. However, our bodies cannot tell the difference between normal iodine and radioactive iodine. To prevent the people in Japan from absorbing radioactive iodine into their thyroid glands, they were issued with iodine tablets (Figure 3.5.7). Taking an iodine tablet each day ensured that the thyroid gland was saturated with iodine and so any radioactive iodine ingested by eating contaminated food would not be taken into the body and deposited in the thyroid.

FIGURE 3.5.7 Japanese children in Japan receiving iodine tablets to prevent their body ingesting radioactive iodine after the Fukushima nuclear accident in 2011.

Many victims of the Chernobyl nuclear disaster in 1986 died of thyroid cancer years after the accident. They ingested radioactive iodine and this accumulated in the thyroid gland, eventually leading to cancer.

Typical exposures to the targeted area for various medical procedures are given in Table 3.5.4.

TABLE 3.5.4 Typical radiation exposures for various medical procedures.

Medical exposures	30 μSv for a chest X-ray 300 μSv for a pelvic X-ray 5000 μSv for a CT scan 40 000 000 μSv for a course of radiotherapy using cobalt-60

Detecting cancer with radioactive tracers

Cancers that form on the skin can often be detected by a simple external examination. However, in order to diagnose the presence of cancerous growths at specific sites inside the body, a variety of radioisotopes tagged to particular drugs are used. The radioisotope is known as a radioactive tracer. These drugs, radiopharmaceuticals, can be administered by swallowing (ingestion), inhalation or injection. In Figure 3.5.9 a gamma-ray camera is being used to perform a bone scan. This patient has been injected with the radioisotope technetium-99m. This isotope is a gamma emitter with a half-life of 6 hours. The camera detects the emitted gamma-rays and produces an image on a computer screen.

FIGURE 3.5.9 A cancer patient receiving a bone scan with a gamma-ray camera.

The radioisotope used in the radiopharmaceutical depends on the site of the suspected tumour. The body naturally distributes different elements to different organs. For example, iodine is sent to the thyroid gland by the liver. So if a radiopharmaceutical containing radioactive iodine is ingested, most of this iodine will end up in the thyroid.

When the tracer has reached the target organ, a radiation scan is taken with a gamma-ray camera. An unusual pattern on the scan indicates a possible cancerous tumour. The radioisotopes used for this type of diagnosis need to be gamma-ray emitters so that the radiation has enough penetrating ability to pass out of the body to reach the detector—the gamma-ray camera. The isotope should have a relatively short half-life so that the patient is not subjected to any unnecessary long-term exposure to radiation.

The most commonly used radioactive tracer is technetium-99m. It is produced on site at hospitals with small nuclear generators. Technetium-99m is a gamma emitter with a half-life of 6 hours and is used to monitor the state of many organs in the body.

Radioactive tracers are also used to monitor other bodily functions. Some examples are shown in Table 3.5.5.

TABLE 3.5.5 Some radioactive tracers and their target organs.

Radioactive tracer	Function monitored
iodine-123	function of thyroid gland
xenon-133	function of lungs
phosphorus-32	blood flow through body
iron-59	level of iron uptake by spleen
technetium-99m	blood flow in brain, lungs and heart function of liver metabolism of bones

PHYSICS IN ACTION

Scintigraphy

The radioactive isotopes of an element have chemical properties that are identical to those of the non-radioactive isotopes. This feature is used by applying radioactive isotopes of elements of different compounds and, in this way, obtaining **tracers**.

Due to the fact that a tracer continuously emits radiation, it is possible to observe what happens to it, to see the path along which it is moving, and the reactions in which it becomes involved.

The tracer method is used, for example, in medicine (e.g., in a scintigraphic examination) and in industry (e.g., examination of the water-tightness of a water system). Figure 3.5.10 shows a scintigram of the human central nervous system (CNS), showing the brain at the top with the spinal cord running below it. The image was obtained by injecting a radioactive tracer into the subject which concentrates in the tissues of the CNS. Gamma rays emitted from the tracer are resolved by a gamma camera's crystal scintillator as flashes of light and are processed electronically to give a map of distribution of radioactivity in the CNS.

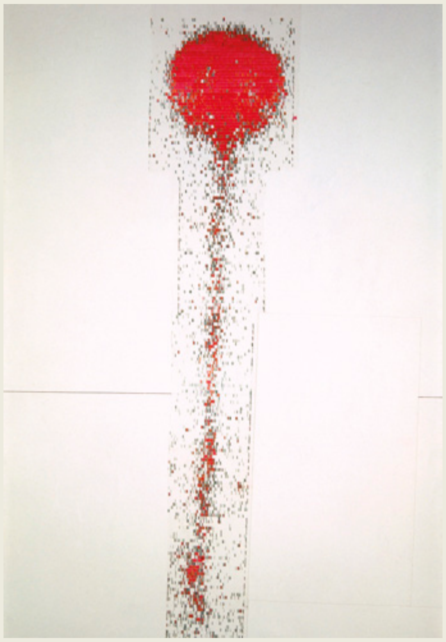


FIGURE 3.5.10 Scintigram of the human central nervous system.

3.5 Review

SUMMARY

- Alpha, beta, gamma and high-energy electromagnetic radiation are ionising and are harmful to biological tissues at high levels.
- Exposure to background ionising radiation is a natural part of our existence.
- Unnecessary exposure to high levels of ionising radiation can be dangerous and should be avoided.
- Absorbed dose (AD) is a measure of the radiation dose per kilogram of irradiated tissue. AD is measured in grays (Gy).

$$AD = \frac{\text{energy absorbed by tissue (J)}}{\text{mass of tissue (kg)}}.$$

- The quality factor (QF) of radiation is a number that indicates the relative damaging effect of the particular radiation.
- Dose equivalent (DE) gives a measure of the biological damage that a dose of radiation causes. DE = AD × QF and is measured in sieverts (Sv).
- High doses of targeted radiation can be used to destroy cancerous cells.
- Radiation from radioisotopes can be used as a diagnostic tool in medicine and industry.

KEY QUESTIONS

1 Which of the following is the most dangerous exposure to radiation?
- **A** 1 Gy of alpha radiation
- **B** 1 Gy of beta radiation
- **C** 1 Gy of gamma radiation
- **D** All are equally damaging

2 A geologist receives a dose of 250 μSv. Which of the following is the most damaging dose?
- **A** 250 μSv of alpha radiation
- **B** 250 μSv of beta radiation
- **C** 250 μSv of gamma radiation
- **D** All doses are equally damaging

3 An 80 kg tourist absorbs a gamma radiation dose of 200 μGy on a return flight from London.
- **a** Calculate the dose equivalent received by the tourist.
- **b** Calculate the amount of radiation energy that has been absorbed by the tourist.

4 Rank the following radiation doses from the most damaging to the least damaging.
- **A** 250 μGy of gamma radiation.
- **B** 20 μGy of alpha radiation.
- **C** 50 μGy of beta radiation.
- **D** 30 μGy of neutron radiation.

5 When in space, astronauts receive a radiation dose of about 1000 μSv per day. The normal annual background dose on Earth is 2 mSv.
- **a** How many days does it take for astronauts to exceed the normal background dose?
- **b** The record for time spent in space is 879 days, held by cosmonaut Gennady Padalka. How much radiation in millisieverts (mSv) was he exposed to in this time?

6 To treat cancer of the uterus, a radioactive source is implanted directly into the affected region. If the uterus receives a dose of 0.40 Gy h^{-1} from the source, how many hours should it be left there to deliver a dose of 36 Gy?

7 Which of the following is the most appropriate for use as a radioactive tracer so that a brain tumour can be detected by a gamma camera?
- **A** radon-222, α emitter, half-life = 3.8 days
- **B** sulfur-35, β emitter, half-life = 97 days
- **C** cobalt-60, γ emitter, half-life = 5.3 years
- **D** technetium-99m, γ emitter, half-life = 6 hours

Chapter review

03

KEY TERMS

absorbed dose
activity
alpha particle
antineutrino
artificial transmutation
atomic number
background radiation
beta particle
daughter nucleus
decay series
deuterium
dose equivalent
electromagnetic radiation
electron
electronvolt
electrostatic force
gamma ray
Geiger counter
half-life
ionising ability
ionising radiation
isotope
mass number
neutral
neutron
non-ionising radiation
nuclear transmutation
nucleon
nucleus
nuclide
parent nucleus
penetrating ability
positron
proton
quality factor
radiation
radioactive
radioisotope
spontaneous transmutation
strong nuclear force
tracer
tritium

1 How many protons and neutrons are in the $^{45}_{20}\text{Ca}$ nuclide?

2 Use the periodic table in Figure 3.1.7 on page 49 to determine the number of protons, neutrons and nucleons in cobalt-60.

3 Determine the nature of the unknown, X, for the following transmutation:

$^{60m}_{27}\text{Co} \rightarrow {}^{60}_{27}\text{Co} + X$

(60m means the nuclide is metastable and has a higher level of stability than very short-lived isotopes. The mass number is still 60.)

4 What type of radiation does potassium-48 (atomic number 19) emit? Use Figure 3.1.10 on page 51 to answer this question.

5 Identify each of these radiation types:

a $^{0}_{-1}A$

b $^{1}_{1}B$

c $^{4}_{2}C$

d $^{1}_{0}D$

e $^{0}_{0}E$

f $^{0}_{1}F$

6 Some nuclei can be made unstable by firing neutrons into them. The neutron is captured and the nucleus becomes unstable. The nuclear equation when the stable isotope boron-10 transmutates by neutron capture into a different element, X, by emitting alpha particles is:

$^{10}_{5}\text{B} + {}^{1}_{0}\text{n} \rightarrow X + {}^{4}_{2}\text{He}$

Identify the unknown element, X, and give its atomic and mass numbers.

7 Identify each of the unknown particles X and Y in the following nuclear transmutations.

a $^{14}_{7}\text{N} + \alpha \rightarrow {}^{17}_{8}\text{O} + X$

b $^{27}_{13}\text{Al} + Y \rightarrow {}^{27}_{12}\text{Mg} + {}^{1}_{1}\text{H}$

8 Find the values of x and y in each of these radioactive decay equations.

a $^{208}_{81}\text{Ti} \rightarrow {}^{x}_{y}\text{Pb} + \beta^{-}$

b $^{180}_{80}\text{Hg} \rightarrow {}^{x}_{y}\text{Pt} + \alpha$

9 Fluorine-18 is a radioisotope that is used for detecting tumours. It is formed when radioactive neon-18 decays by positron emission. Fluorine-18 in turn also decays by positron emission. The equations are as follows:

$^{18}_{10}\text{Ne} \rightarrow {}^{a}_{b}\text{X} + {}^{0}_{+1}\beta$

$^{a}_{b}X \rightarrow {}^{c}_{d}Y + {}^{0}_{+1}\beta$

Determine the values of a, b, c, d and identify X and Y, which are the daughter nuclei that result from this process.

10 The radioisotope nitrogen-12 decays by emitting a positron and a neutrino. The decay equation for nitrogen-12 is:

$^{12}_{7}\text{N} \rightarrow X + {}^{0}_{+1}\beta + {}^{0}_{0}\nu$

Identify X.

11 A stable isotope of neon has 10 protons and 10 neutrons in each nucleus. Every proton is repelling all the other protons. Why is the nucleus stable?

12 Which type of radiation out of alpha, beta, and gamma:

a is the fastest?

b has the greatest penetrating power?

13 Health workers who deal with radiation to treat cancer must wear a lead vest to protect their vital organs from exposure. Which type(s) of radiation is the lead apron shielding them from?

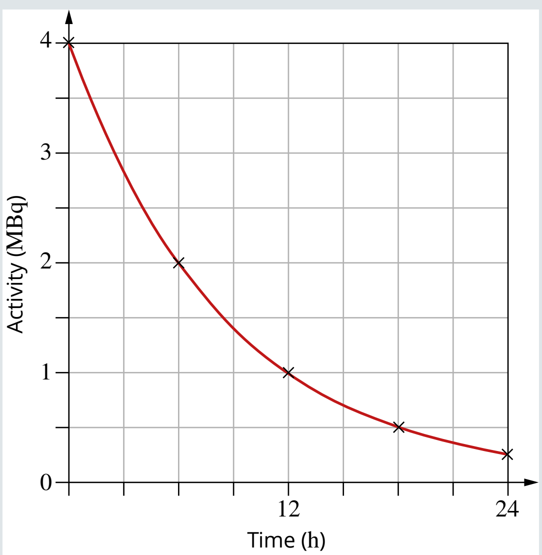
14 A nuclear physicist was bombarding a sample of beryllium-7 with a beam of electrons in an effort to smash the electrons into the beryllium nuclei. Why would it be quite difficult for a collision between the electrons and the nuclei to occur?

15 A radioactive isotope *X* has a half-life of 20 minutes. A sample starts with 6.0×10^{14} atoms of the isotope. What amount of the original isotope will remain after 20 minutes?

16 Radioisotope *Y* has a half-life of 3.0 hours. A sample starts with 5.6×10^{15} atoms of the radioisotope. How many atoms of *Y* remain after 9.0 hours?

17 Uranium-235 has a half-life of 700 000 000 years (700 million years), while uranium-238 has a half-life many times longer of 4.5×10^9 years. In samples of 1 kg of each of these pure radioisotopes, which one would have the greater activity?

18 The decay curve for a sample of the radioisotope technetium-99m is shown. It has an initial activity of 4.0×10^6 Bq.



a What is the activity of this sample after one half-life?

b From the graph, what is the half-life of technetium-99m?

c If the sample is produced in a hospital at 4 pm, what is its activity when it is used at 10 am the next day?

19 Protactinium-234 is a radioactive element with a half-life of 70 s. If a sample of this radioisotope contains 6.0×10^{10} nuclei, how many nuclei of this element will remain after 140 s?

20 Radiotherapy treatment of brain tumours involves irradiating the target area with radiation from an external source. Why is cobalt-60 (gamma emitter with a half-life of 5.3 years) generally used as the radiation source for this treatment?

21 An airline pilot of mass 90 kg absorbs a gamma radiation dose of 300 mGy during a return flight to New York. Calculate the dose equivalent that has been received in mSv.

22 In a major incident in a nuclear reactor, a 75 kg employee received a full-body absorbed radiation dose of 5.0 Gy. The radiation was gamma rays.

a Calculate the amount of energy that was absorbed during this exposure. Give your answer to 2 significant figures.

b Calculate the dose equivalent for this person. Give your answer to 2 significant figures.

23 A worker in an X-ray clinic takes an average of 10 X-ray photographs each working day and receives an annual radiation dose equivalent of 7900 μSv.

a Calculate the dose, in μSv, that the worker receives from each X-ray photograph. (Assume they work for 5 days per week for 45 weeks a year, and give your answer to 1 significant figure.)

b How many times greater than the normal background radiation dose is the worker's annual dose?

CHAPTER

04 Fission and fusion

This chapter looks at typical nuclear fission and fusion reactions, the forces that act within the nucleus, energy transfer and important transformation phenomena in stars and in the production of nuclear energy. It also examines the benefits and risks of using nuclear power as an energy source for society.

Science as a Human Endeavour

Qualitative and quantitative analyses of relative risk are used to inform community debates about the use of radioactive materials and nuclear reactions for a range of applications and purposes, including:

- A fission chain reaction is a self-sustaining process that may be controlled to produce thermal energy, or uncontrolled to release energy explosively if its critical mass is exceeded.
- nuclear power stations employ a variety of safety mechanisms to prevent nuclear accidents, including shielding, moderators, cooling systems, and radiation monitors
- the management of nuclear waste is based on the knowledge of the behaviour of radiation.

Science Understanding

- neutron-induced nuclear fission is a reaction in which a heavy nuclide captures a neutron and then splits into smaller radioactive nuclides with the release of energy
- Einstein's mass/energy relationship relates the binding energy of a nucleus to its mass defect

  This includes applying the relationship

  $\Delta E = \Delta mc^2$
- Einstein's mass/energy relationship also applies to all energy changes and enables the energy released in nuclear reactions to be determined from the mass change in the reaction

  This includes applying the relationship

  $\Delta E = \Delta mc^2$
- nuclear fusion is a reaction in which light nuclides combine to form a heavier nuclide, with the release of energy
- more energy is released per nucleon in nuclear fusion than in nuclear fission because a greater percentage of the mass is transformed into energy

4.1 Nuclear fission and energy

In 1905, Albert Einstein theorised that mass, m, and energy, E, are equivalent through the equation $E = mc^2$. This led to the realisation that vast amounts of energy lie unharnessed within the nuclei of atoms. The ramifications of Einstein's work and the discovery of nuclear fission were realised in 1945 with the explosion of the first atomic bomb in the desert near Alamogordo in New Mexico, USA (Figure 4.1.1). In this section, nuclear fission and the energy that it can unleash will be explored.

FIGURE 4.1.1 An atomic bomb explosion and its associated mushroom cloud.

INSIDE THE NUCLEUS

The current understanding of the basic properties and structure of the nucleus is the result of intense scientific investigation in the early part of the twentieth century. Physicists such as Becquerel, Rutherford, Chadwick, Geiger, Marsden and Harkins were instrumental in the development of the model of the nucleus that exists today. These renowned scientists are shown in Figure 4.1.2.

Recall from Chapter 3 that there is a strong nuclear force that acts within the nucleus to overcome the electrostatic repulsion of the protons. This force holds the nucleus together.

FIGURE 4.1.2 (a) Henri Becquerel, (b) Ernest Rutherford, (c) James Chadwick, (d) Hans Geiger, (e) Ernest Marsden and (f) William Harkins.

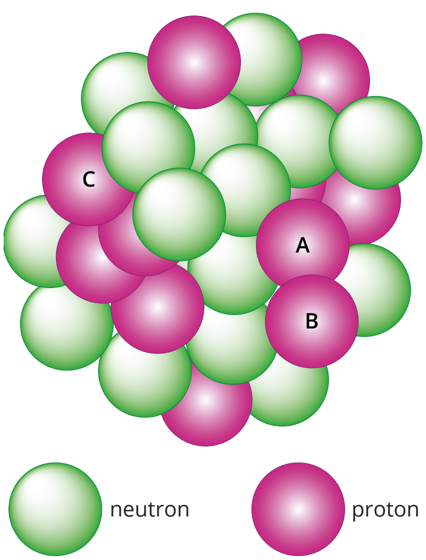


FIGURE 4.1.3 The interaction between the electrostatic and strong nuclear forces acting in a nucleus.

An example of how the electrostatic and strong nuclear forces act in a nucleus is shown in Figure 4.1.3. In this example, proton A both attracts and repels proton B, but, at short distances, the attraction due to the strong nuclear force is much greater than the repulsion due to the electrostatic force. Proton A also both attracts and repels proton C, but because of the greater distance between them, the force of repulsion is larger. However, proton A and proton C do not fly apart due to the strong attractive forces exerted on them by adjacent neutrons.

NUCLEAR FISSION

The discovery of the neutron by James Chadwick in 1932 enabled scientists to explore the behaviour of larger atomic nuclei. Up until then, physicists such as Enrico Fermi had been firing alpha-particles at target nuclei and analysing the results. Chadwick found that with larger target nuclei, the positive alpha-particles were too strongly repelled from the positively charged nuclei and collisions did not occur.

The advantage of a neutron is that it is neutral and so is not repelled by any target nucleus. The bombarding neutrons can be absorbed into the nucleus of the target atom, as shown in Figure 4.1.4. This makes neutrons very useful as a form of radiation. They are used in many experiments to artificially transmutate different isotopes.

PHYSICSFILE

Strong nuclear force

The existence of the strong nuclear force was first proposed by Japanese theoretical physicist Hideki Yukawa in 1935. However, the properties of this force are so complex that it took until 1975 for physicists to develop a mathematical model that could successfully describe it.

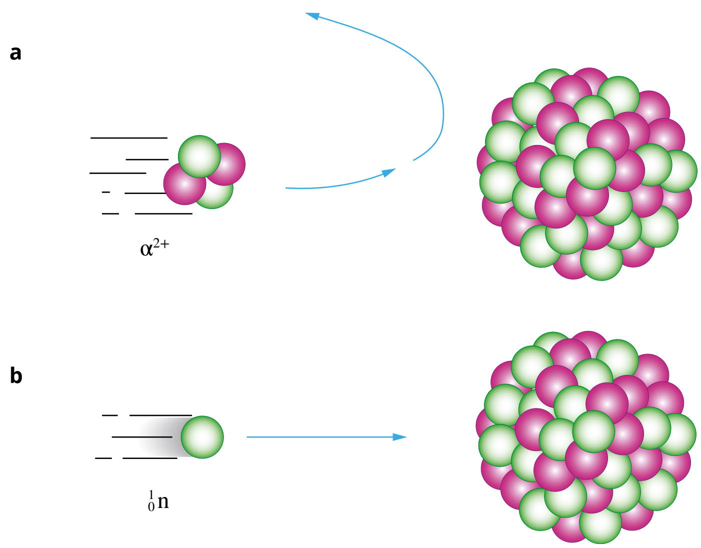


FIGURE 4.1.4 (a) Charged α-particles are repelled by a nucleus. (b) Uncharged neutrons are able to smash into a nucleus.

Nuclear **fission** occurs when an atomic nucleus splits into two or more pieces. This is usually triggered or induced by the absorption of a neutron, as shown in Figure 4.1.5. Nuclides that are capable of undergoing nuclear fission after absorbing a neutron are said to be **fissile**. Fissile nuclides are all elements with high atomic numbers, very few of which exist in nature.

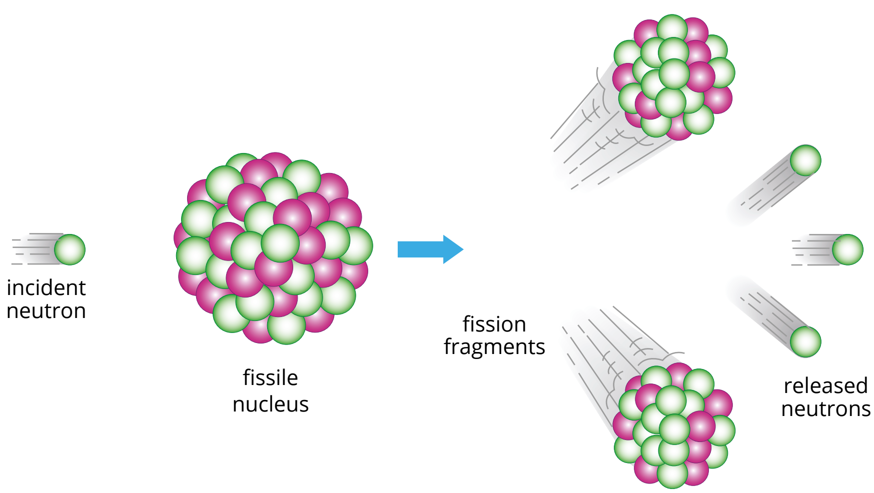


FIGURE 4.1.5 Nuclear fission is the splitting of a nucleus.

Uranium-235 and plutonium-239 are fissile and can be made to split when bombarded by a slow-moving neutron. Uranium-238 and thorium-232 require a very high-energy neutron to induce fission, so they are regarded as fissionable, but non-fissile.

Uranium is one of the heaviest naturally occurring elements, has several isotopes and is found in most rocks. It is thought to have formed in supernovas around 6.6 billion years ago. Its isotopes have very long half-lives and the energy generated from their radioactive decay is thought to be the main source of heat to the Earth's core, resulting in convection currents and continental drift.

RELEASE OF NEUTRONS DURING FISSION

Uranium-235 and plutonium-239 are the fissile nuclides most commonly used in nuclear reactors and nuclear weapons. They are more fissile than uranium-238 and thorium-232.

When a uranium-235 or plutonium-239 nucleus absorbs either a slow- or fast-moving neutron, it becomes unstable and spontaneously undergoes fission. However, fission is more likely to be induced by a slow-moving neutron because it is more easily captured by the target nucleus.

A uranium-235 nucleus may split in many different ways. When a uranium-235 nucleus undergoes fission, it splits into two smaller nuclei plus neutrons. A wide variety of pairs of smaller nuclei are produced, due to the completely random way in which the nucleus splits. Many of the products are themselves radioactive. Figure 4.1.6 shows one outcome but many others are possible. In this example three neutrons are released, along with krypton-91 and barium-142 nuclides. Usually either two or three neutrons are released. For uranium-235, an average of 2.47 neutrons per fission has been determined.

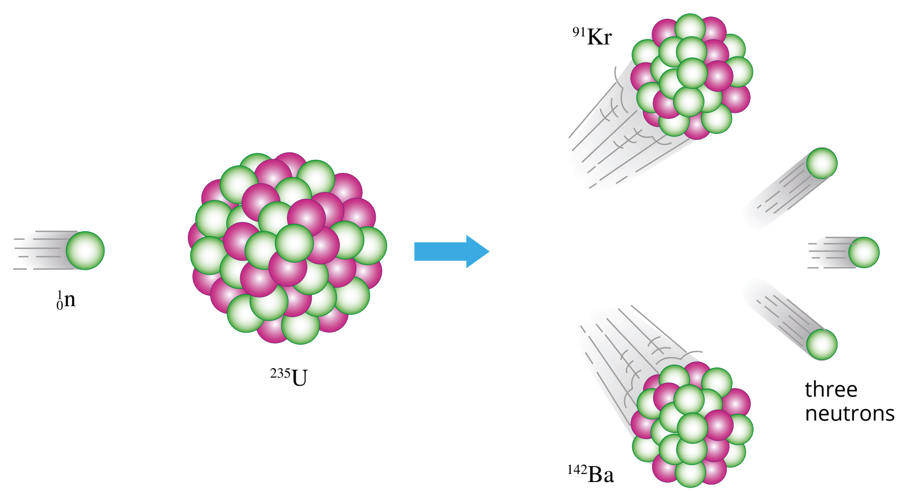


FIGURE 4.1.6 One possible outcome for the neutron-induced fission of uranium-235.

The equation for this reaction is:

$$^{1}_{0}n + ^{235}_{92}U \rightarrow ^{236}_{92}U \rightarrow ^{91}_{36}Kr + ^{142}_{56}Ba + 3^{1}_{0}n + \text{energy}$$

Krypton-91 and barium-142 are known as fission fragments or daughter nuclei. Three neutrons are freed from this uranium nucleus when it splits. Note that in the same way as for radioactive decay, both the atomic number, Z, and mass number, A, are conserved in these nuclear reactions. For the reaction equation shown, the atomic numbers on either side of the arrows add up to 92 and the mass numbers add up to 236.

In the end, the decay products of the nuclear fission process form a lethal cocktail of radioactive isotopes. It is these radioactive fission fragments that comprise the bulk of the high-level waste produced by nuclear reactors.

Plutonium-239 will also undergo fission in a variety of ways. It releases an average of 2.89 neutrons per fission, slightly more than uranium-235, but it does not undergo fission as easily.

HOW TO MEASURE ENERGY IN ELECTRONVOLTS

The energy of moving objects such as cars and tennis balls is measured in joules. However, nuclei, subatomic particles and radioactive emissions have such small amounts of energy that the joule is inappropriate.

The energy of subatomic particles and radiation is usually given in **electronvolts** (eV). One electronvolt is an extremely small amount of energy and is equivalent to 1.60×10^{-19} J.

> An electronvolt is the energy that an electron would gain if it were accelerated by a voltage of 1 volt and is equal to 1.6×10^{-19} J.
>
> To convert from eV to joules: multiply by 1.6×10^{-19} J.
>
> To convert from joules to eV: divide by 1.6×10^{-19} J.

ENERGY RELEASED DURING NUCLEAR REACTIONS

It is well established that the mass of any nucleus is always less than the mass of its individual nucleons. Two separate protons and two separate neutrons will have slightly more total mass than a helium nucleus.

Albert Einstein, pictured in Figure 4.1.7, provided the explanation of the origins of this missing mass. He showed that *mass* and *energy* were not completely independent quantities. Indeed, mass can be converted into energy and energy can be converted into mass.

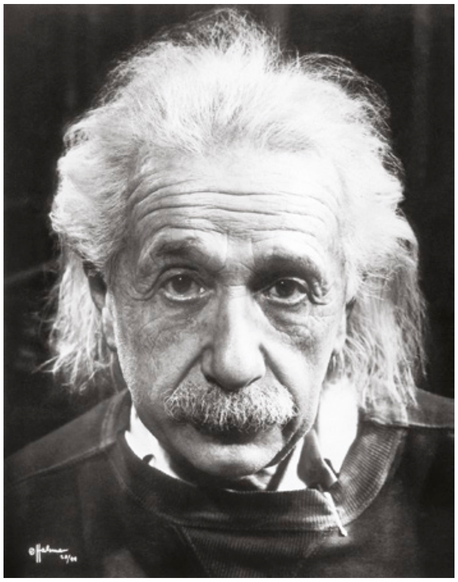


FIGURE 4.1.7 Albert Einstein.

If you wanted to separate a helium nucleus into four free nucleons, you would need to add energy to the nucleus. This energy is the **binding energy** of the nucleus. The free nucleons will have more energy and so, according to Einstein, will have greater mass.

The energy released as a result of a mass defect (mass decrease) is given by Einstein's famous equation:

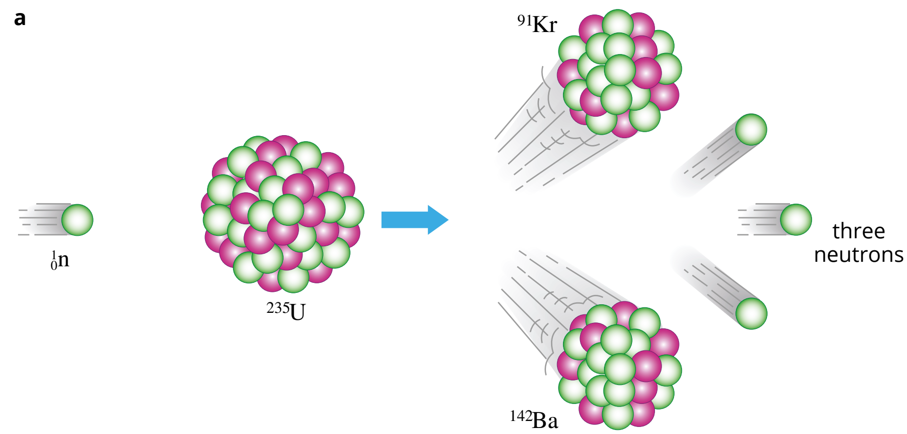
> $\Delta E = \Delta mc^2$
>
> where ΔE is energy (J)
>
> Δm is the mass defect (the decrease in mass, in kg)
>
> c is the speed of light $= 3.0 \times 10^8$ m s^{-1}

The chemical reactions that you have probably performed at school typically release only a few electronvolts of energy. Compared with this, an enormous amount of energy is released during nuclear reactions. This has made nuclear energy a major focus of scientific research over the past century.

During radioactive decay millions of electronvolts of energy can be released. Alpha particle decay usually involves the release of 5–10 MeV (5–10 million electronvolts) of energy. Nuclear fission involves much more energy again, typically around 200 MeV. This energy is mainly in the form of the kinetic energy of the fission fragments and neutrons, as well as the emission of energy as gamma radiation.

During any fission reaction, the combined mass of the incident neutron and the target nucleus is always slightly greater than the combined mass of the fission fragments and the released neutrons. For example, in Figure 4.1.8 on page 92, the mass of the incident neutron and the uranium-235 nucleus is greater than the combined masses of the fission products—barium-142, krypton-91 and the three neutrons. This missing mass is converted into energy according to the equation $\Delta E = \Delta mc^2$. In this case 200 MeV of energy is released.

Only a very small proportion of the original mass of the nuclei is available as usable energy—typically around 0.1%. If you had a 1.000 kg block of pure uranium-235 that underwent fission completely, at the end you would have a block of radioactive fission fragments with a mass of around 0.999 kg.



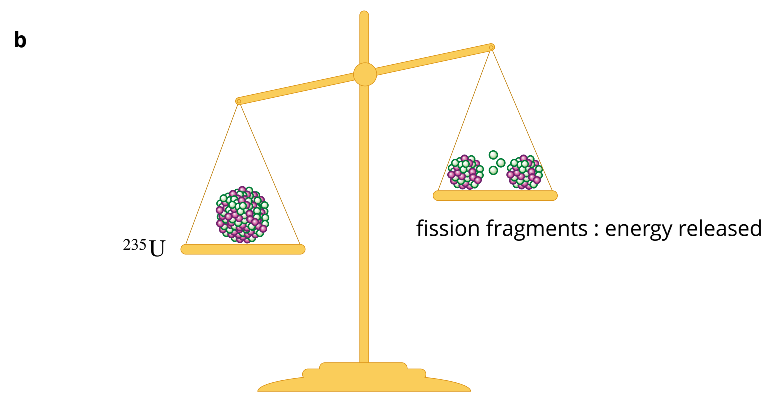


FIGURE 4.1.8 The mass of fission products is less than the mass of the incident neutron and target atom.

Worked example 4.1.1

FISSION

Plutonium-239 is a fissile material. When a plutonium-239 nucleus is struck by and absorbs a neutron, it can split in many different ways. Consider the example of a nucleus that splits into barium-145 and strontium-93 and releases some neutrons.

The nuclear equation for this is:

$$^1_0n + ^{239}_{94}Pu \rightarrow ^{145}_{56}Ba + ^{93}_{38}Sr + a^1_0n + \text{energy}$$

a How many neutrons are released during this fission process, i.e. what is the value of *a*?	
Thinking	**Working**
Analyse the mass numbers (*A*).	$1 + 239 = 145 + 93 + (a \times 1)$ $a = (1 + 239) - (145 + 93)$ $= 2$ 2 neutrons are released during this fission.

b During this single fission reaction, there is a loss of mass (a mass defect) of 3.07×10^{-28} kg. Calculate the amount of energy that is released during the fission of a single plutonium-239 nucleus. Answer in both MeV and joules.	
Thinking	**Working**
The energy released during the fission of this plutonium nucleus can be found by using $E = mc^2$.	$\Delta E = \Delta mc^2$ $= (3.07 \times 10^{-28}) \times (3.00 \times 10^8)^2$ $= 2.76 \times 10^{-11}$ J
To convert J into eV, divide by 1.6×10^{-19}. Remember that 1 MeV = 10^6 eV.	$E = \frac{2.76 \times 10^{-11}}{1.6 \times 10^{-19}}$ $= 1.73 \times 10^8$ eV $= 173$ MeV

c The combined mass of the plutonium nucleus and bombarding neutron is 3.99×10^{-25} kg. What percentage of this initial mass is converted into the energy produced during the fission process?	
Thinking	**Working**
Use the relationship percentage of initial mass converted into energy $= \frac{\text{mass defect}}{\text{initial mass}} \times \frac{100}{1}$	percentage of initial mass converted into energy $= \frac{\text{mass defect}}{\text{initial mass}} \times \frac{100}{1}$ $= \frac{3.07 \times 10^{-28}}{3.99 \times 10^{-25}} \times \frac{100}{1}$ $= 0.0769\%$ This is a very small percentage loss in mass.

Worked example: Try yourself 4.1.1

FISSION

Plutonium-239 is a fissile material. When a plutonium-239 nucleus is struck by and absorbs a neutron, it can split in many different ways. Consider the example of a nucleus that splits into lanthanum-143 and rubidium-94 and releases some neutrons.

The nuclear equation for this is:

$$^{1}_{0}n + ^{239}_{94}Pu \rightarrow ^{143}_{57}La + ^{94}_{37}Rb + a^{1}_{0}n + \text{energy}$$

a How many neutrons are released during this fission process, i.e. what is the value of *a*?

b During this single fission reaction, there is a loss of mass (a mass defect) of 4.58×10^{-28} kg. Calculate the amount of energy that is released during fission of a single plutonium-239 nucleus. Give your answer in both MeV and joules to two significant figures.

c The combined mass of the plutonium nucleus and bombarding neutron is 2.86×10^{-25} kg. What percentage of this initial mass is converted into the energy produced during the fission process?

PHYSICS IN ACTION

Enrico Fermi, Lise Meitner and nuclear fission

Enrico Fermi, pictured in Figure 4.1.9, was born in Italy in 1901. He completed his doctorate and post-doctorate work in physics at the University of Pisa and in Germany. Fermi had emigrated to the USA by the time the nuclear age dawned in the 1930s. The neutron had just been discovered in 1932, which enabled scientists to fire neutral particles at atomic nuclei for the first time. Fermi was at the forefront of this research.

Fermi bombarded uranium-238 atoms with neutrons and found that uranium-238 nuclei absorbed the neutrons and formed a radioactive isotope of uranium. This isotope then decayed by emitting a beta-minus particle to become neptunium, which then emitted another beta-minus particle to become plutonium, two completely undiscovered elements. Fermi had successfully produced the world's first artificial and **transuranic** (i.e. after uranium) elements. The nuclear reactions for this process are:

$$^{1}_{0}n + ^{238}_{92}U \rightarrow ^{239}_{92}U$$

$$^{239}_{92}U \rightarrow ^{238}_{93}Np + ^{0}_{-1}\beta$$

$$^{239}_{93}Np \rightarrow ^{239}_{94}Pu + ^{0}_{-1}\beta$$

In 1938, following on from Fermi's work, two German scientists, Otto Hahn and Fritz Strassmann, were also bombarding uranium ($Z = 92$) in an attempt to produce some transuranic elements ($Z > 92$). They found that, rather than producing larger elements, they were getting isotopes of barium ($Z = 56$). Hahn wrote to his colleague Lise Meitner, pictured in Figure 4.1.10, about this unexpected result. She then discussed this with her nephew Otto Frisch, a nuclear physicist, and realised that the bombarding neutrons were causing the uranium nuclei to split.

If barium ($Z = 56$) was one of the products, then krypton ($Z = 36$) must be another. This was found to be the case. It was Frisch who coined the term 'fission' and Meitner who proposed that energy would be released during this process.

After the start of World War II, Enrico Fermi was commissioned by President Roosevelt to design and build a device that would sustain the fission process in the form of a chain reaction. In 1942, Fermi succeeded in this task. A squash court at the University of Chicago was used as the site for the world's first nuclear reactor. It produced less than 1 W of power—not even enough to power a small light globe! This sounds like a bit of a failure, but in fact, achieving fission for the first time was a very important breakthrough. The reactor was later modified to produce about 200 W. Fermi died of cancer in 1954. One year after his death, the element with atomic number 100 was artificially produced and named fermium, Fm, in his honour.

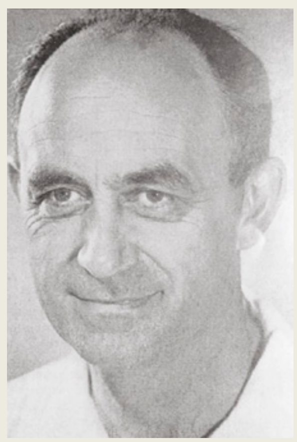


FIGURE 4.1.9 Enrico Fermi.

FIGURE 4.1.10 Lise Meitner.

4.1 Review

SUMMARY

- Within a nucleus, forces of attraction and repulsion are acting. The long-distance electrostatic force of repulsion acts between the protons. A short-distance strong nuclear force of attraction acts between every nucleon.
- Nuclear fission occurs when a nucleus is made to split and release a number of neutrons. This can be induced by striking a fissile nucleus with a neutron. A relatively large amount of energy is released during this process.
- When fission occurs, the mass of the fission fragments is always less than the mass of the original particles. This decrease in mass is proportional to energy released, as given by $\Delta E = \Delta mc^2$.
- The energy of subatomic particles can be measured in joules but is usually measured in electronvolts (eV).
- $1\text{ eV} = 1.6 \times 10^{-19}\text{ J}$
  - To convert from eV to joules: multiply by 1.6×10^{-19}.
  - To convert from joules to eV: divide by 1.6×10^{-19}.

KEY QUESTIONS

1 What is the strong nuclear force?

2 Why is dangerous waste produced by nuclear fission reactors?

3 Consider one particular neutron in the nucleus of a gold atom. Describe the forces that the neutron experiences from other nucleons.

4 Convert 5.0 MeV into joules.

5 Convert 6.0×10^{-15} J into eV.

6 Which of the nuclides below are fissile and which are non-fissile?
cobalt-60, uranium-235, uranium-238, plutonium-239

7 Determine the number of neutrons (x) released in this fission reaction:

$$ {}^{1}_{0}\text{n} + {}^{235}_{92}\text{U} \rightarrow {}^{148}_{57}\text{La} + {}^{85}_{35}\text{Br} + x{}^{1}_{0}\text{n} $$

8 The mass defect during the process in question 7 is 2.12×10^{-28} kg.

a Calculate the energy (in joules) released per uranium-235 atom in this fission process.

b Express this energy in electronvolts.

9 Einstein said that mass and energy are equivalent. In one particular nuclear reaction there is a decrease in mass of 3.48×10^{-28} kg. What energy does this represent in both joules and electron volts?

10 During the radioactive decay of an excited cobalt-60 nucleus to a stable cobalt-60 nucleus a gamma ray of energy 1.33 MeV is released. Calculate the mass defect (in kg) of this cobalt-60 nucleus.

11 Determine the value of the unknown mass number x and atomic number y in the fission reaction below.

$$ {}^{1}_{0}\text{n} + {}^{x}_{94}\text{Pu} \rightarrow {}^{130}_{54}\text{Xe} + {}^{106}_{y}\text{Xr} + 4{}^{1}_{0}\text{n} $$

4.2 Chain reactions and nuclear reactors

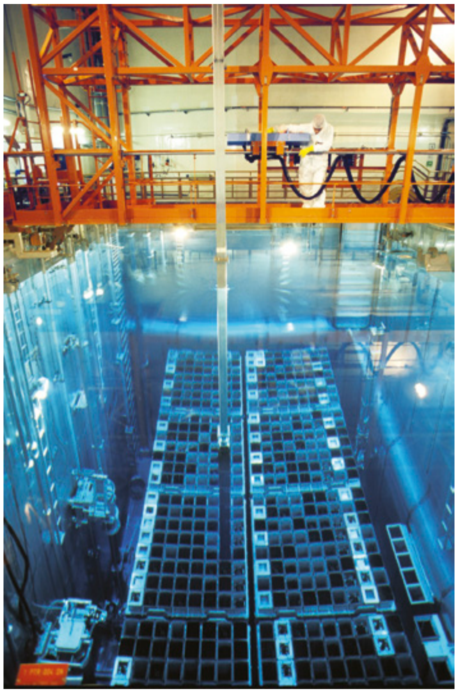


FIGURE 4.2.1 Spent fuel rods in a cooling pond at a French nuclear power station.

In the years following the discovery of nuclear fission, intensive scientific and military research was undertaken into the use of fission for nuclear weapons and for generating electricity. Nuclear fission weapons were developed in World War II and are capable of causing death and devastation on a massive scale. This was tragically evident in the bombing of Hiroshima and Nagasaki in 1945. In this section, you will learn about the work scientists undertook to go from the fission of a single nucleus to a chain reaction that would release vast amounts of energy, and the types of fuel used in nuclear reactions.

Once it was possible to control nuclear fission within a nuclear reactor it was used as a way of producing electricity. In this section you will also learn how nuclear fission is harnessed to make nuclear reactors produce electricity.

You will also learn about the nuclear waste produced and how it is managed. Figure 4.2.1 is a photograph of a cooling pond used for storing spent nuclear fuel rods. Over 40 per cent of spent fuel rods produced world-wide are permanently stored in this way. The remainder are reprocessed and the fissile material is reused as nuclear fuel, the fission products being stored permanently in underground bunkers.

CHAIN REACTIONS AND CRITICAL MASS

Chain reactions

During World War II, the United States and Great Britain established a massive top-secret research project, named the Manhattan Project, to design and build the first atomic bombs. The scientists working on this project knew that nuclear energy could be released from a single fissile nucleus, but the problem they faced was how to obtain this energy from a vast number of fissile nuclei and hence create an explosive device.

When uranium-235 undergoes fission, it releases two or three neutrons each time. Each of these neutrons is then able to cause fission in another uranium-235 nucleus, which in turn will also release another two or three neutrons. Within a very short time, the number of released neutrons and fission reactions escalates in a process known as a **chain reaction**.

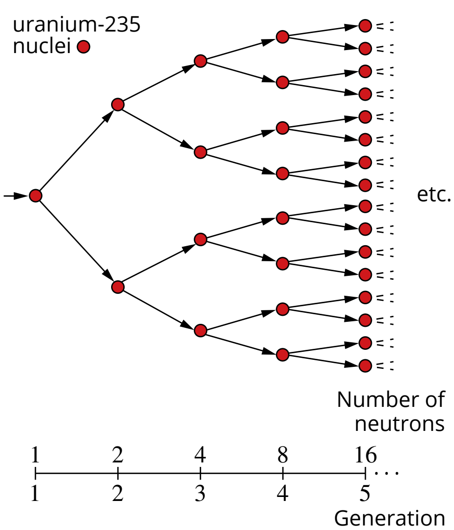


FIGURE 4.2.2 A chain reaction of nuclei.

Figure 4.2.2 shows two neutrons being released during the fission reaction. Each of these neutrons in turn causes fission, releasing yet more neutrons. The number of nuclei undergoing fission doubles with each generation, and after five nuclear generations, 16 neutrons are produced. Within a small fraction of a second an enormous number of nuclei have undergone fission. Only a miniscule amount of energy (of the order of 10^{-13} J) is released by each fission reaction, but in this uncontrolled chain reaction there are so many fission reactions occurring in such a short time that a massive explosion results. In 1 kg of uranium-235, so many reactions occur that about 8 × 1013 J of energy is released in just over one-millionth of a second.

Nuclear fuel

The Earth contains many naturally occurring radioisotopes. In the 4.5 billion years since the Earth was formed these have been decaying to form more stable isotopes. Uranium is one such element. Its two most common isotopes are uranium-238 and uranium-235. These two isotopes have half-lives of 4.5 billion years and 710 million years respectively, and so uranium-235 has been decaying at a faster rate than uranium-238. This means that far less uranium-235 remains in the Earth's crust, so that today the uranium that is mined from the ground consists of:

- 99.3% uranium-238—the non-fissile isotope
- 0.7% uranium-235—the readily fissile isotope.

This means that a chain reaction cannot occur in a sample of uranium taken from the ground because the proportion of fissile uranium-235 is far too low. To be useful as a nuclear fuel, the ore has to be enriched. This involves increasing the proportion of uranium-235 relative to uranium-238, a very difficult and expensive process. The slightly different masses of these two isotopes enables them to be separated. The three common enrichment methods are the ultracentrifuge and electromagnetic and gaseous diffusion separation techniques.

Nuclear weapons require fissile material that has been enriched to over 90% purity. The bomb that was dropped over Hiroshima contained 40 kg of 95% pure uranium-235. Nuclear reactors require fissile material enriched to only about 4% uranium-235.

Critical mass

In their efforts to produce an explosion, the scientists working on the Manhattan Project had to establish a nuclear chain reaction in a sample of nuclear fuel. They found that the explosive ability of a sample of fissile material depended on its purity, shape and size.

In a sample of nuclear material where the concentration of uranium-235 or plutonium-239 is too low, a chain reaction cannot be established. This is because the neutrons have only a small chance of being absorbed by fissile nuclei and causing a further fission reaction. The chain reaction will die out. The fuel used in nuclear fission weapons is enriched to a high degree of purity so that a chain reaction can be easily sustained.

The shape of the nuclear fuel is an important factor in its explosive ability. A 20 kg sample of enriched uranium-235 in the shape of a sphere will spontaneously explode, whereas 20 kg of uranium-235 flattened into a sheet will not. The flat piece has a very large surface area and so an enormous number of neutrons are able to escape from the uranium into the air. These neutrons do not cause further fission reactions and so the chain reaction will die out. In the spherical piece of uranium, the surface area is much smaller and a greater proportion of neutrons remain in the uranium to sustain the chain reaction. Figure 4.2.3 shows how a greater proportion of neutrons are retained within a spherical shape than in a flat shape.

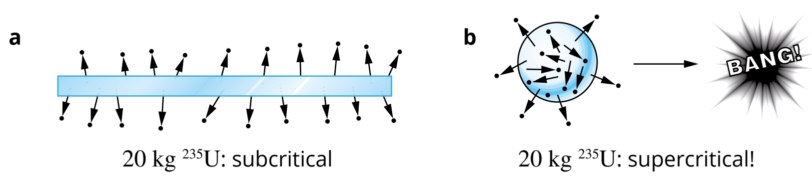


FIGURE 4.2.3 (a) A large proportion of neutrons escape the flat piece of uranium. (b) A sufficient proportion of neutrons remain to maintain the chain reaction—leading to an explosion.

The explosive ability of a fissile material also depends on its physical size. For example, a piece of uranium-235 the size of a marble will not explode but a piece the size of a grapefruit most definitely will. The small piece has more surface area compared with its volume than the large piece. In the marble-sized lump, a greater proportion of neutrons escape into the air and so the chain reaction dies out. This is a **subcritical mass**. In the sample the size of a grapefruit, a higher proportion of neutrons is available to continue the chain reaction within the material. This is a **supercritical mass**, capable of causing a nuclear explosion.

The minimum amount of enriched fissile material in the shape of a sphere that leads to a sustained chain reaction is known as the **critical mass**.

NUCLEAR REACTORS

Since the 1950s, it has been possible to control nuclear fission within a nuclear reactor to produce electrical power for domestic use. Australia does not produce any electricity in this way, but in over 30 countries around the world there are more than 400 nuclear power plants in operation. Many more reactors have been constructed for medical, military and research purposes.

A nuclear power plant generates electricity in much the same way as a coal-burning power plant. The primary difference is how the heat is produced and collected. Most power stations in Australia generate electricity by burning coal, oil or gas to produce heat that creates the steam used to turn the generator turbines. A nuclear power station simply has a different way of producing heat—by nuclear fission.

Thermal nuclear reactor

A thermal nuclear reactor generates energy through the fission of uranium-235, the isotope that is most likely to undergo fission when it is hit by slow moving, or thermal, neutrons.

As they tried to harness the energy from the nuclear fission reactions, nuclear reactor designers had to overcome three major difficulties. The first was that neutrons released from a uranium-235 nucleus undergoing fission are travelling at very high speeds—around 20 000 km s^{-1}. Uranium-235 is most fissile when irradiated by slow-moving neutrons. Thus, these emitted neutrons needed to be slowed down. Second, the fission of each uranium-235 nucleus releases an average of 2.47 neutrons. This can lead to a chain reaction that will result in an explosion. A way had to be found to absorb some of these emitted neutrons and maintain a steady chain reaction. Third, the heat generated in the reactor by the fission process had to be somehow collected and used to create steam to drive the turbine and generate electricity.

There are many different varieties of thermal nuclear reactors, but they all include the following design elements:

- fuel rods—long, thin rods containing pellets of enriched uranium
- a moderator—a material that slows the neutrons
- control rods—material that absorbs neutrons
- a coolant—a liquid to absorb heat energy that has been produced by nuclear fission
- radiation shield—a thick concrete wall that prevents neutrons escaping from the reactor.

Nuclear fuel rods

Uranium-235 is used as nuclear fuel as it is readily fissile with slow-moving neutrons. However, this isotope comprises only 0.7% of naturally occurring uranium, which is not enough to sustain a chain reaction. The predominant (99.3%) isotope, uranium-238, is effectively non-fissile. Therefore, the proportion of uranium-235 in the ore has to be increased. That is, the uranium ore has to be enriched by increasing the uranium-235 content before it can be used as reactor-grade fuel. Uranium enriched for use in a thermal nuclear reactor contains around 96% uranium-238 and 4% uranium-235.

The proportion of uranium-235 is increased to around 4% for fuel rods. The enriched uranium, in pellet form, is then packed into a thin aluminium tube, known as a fuel rod. This is usually 3–5 m long. A large nuclear reactor has over 1000 fuel rods in its core. A fuel rod will eventually become depleted of uranium-235. This means that, over time, the concentration of uranium-235 falls to a level where it cannot sustain the fission chain reaction. Each fuel rod needs to be replaced every 4 years or so, and a typical 1000 MW reactor produces around 25 tonnes of spent fuel each year.

While uranium-238 is not readily fissile, it is classified as 'fertile' because a uranium-238 nucleus can capture a fast neutron from a uranium-235 disintegration and form plutonium-239, which is itself fissile. This reaction yields a similar amount of energy per fission as does uranium-235 and releases sufficient neutrons to sustain a chain reaction.

The initial neutrons given off by the uranium-235 have therefore been able to sustain their own chain reaction while at the same time 'breeding' plutonium-239 to be further used as a fuel. These reactors can be configured to produce more fissionable fuel than they use and are referred to as **fast breeder reactors**.

PHYSICSFILE

The first nuclear reactor

The first nuclear reactor was designed by Enrico Fermi as part of the Manhattan Project during World War II. The reactor contained layers of graphite, and over 40 tonnes of uranium and uranium oxide. It was constructed in a squash court under a grandstand at the University of Chicago. A self-sustained chain reaction was first established at 3:45 pm on 2 December 1942. This historic event marked the beginning of the atomic age. The reactor initially generated less than 1 W of power—not even enough to light a light globe! It was later modified to produce about 200 W.

The moderator

In a nuclear reactor, the problem of fast moving neutrons is overcome by including a material that slows down, or moderates, the speed of the free neutrons. This is known as a moderator. It has been found that substances whose nuclei are small will slow the neutrons down to speeds at which they can be captured by a fissile nucleus. When the emitted neutrons collide with these small nuclei, they lose most of their kinetic energy and so slow down. After many collisions, the neutrons have been slowed down to about $2\,\text{km}\,\text{s}^{-1}$ and have less than 1 eV of energy.

Some materials that are commonly used as moderators are:

- graphite—consisting of carbon atoms
- normal water—H_2O
- heavy water—containing deuterium
- carbon dioxide—CO_2.

Each of these materials works well as a moderator because it slows the neutrons without absorbing a significant number of them. **Heavy water** is the most effective moderator, but is also the most expensive. Water is the cheapest material, but absorbs more neutrons than the others and so reduces the extent of the chain reaction. Graphite is less effective than water because carbon nuclei are heavier than hydrogen nuclei. To lose the same amount of energy, the emitted neutrons have to collide with about 120 carbon nuclei, but only about 25 water molecules.

Control rods

The number of neutrons released during the uncontrolled fission chain reaction grows exponentially, releasing enormous amounts of energy in a split second. A nuclear reactor can also produce great amounts of energy. However, a steady controlled energy release is required. This is achieved by controlling the number of neutrons that are involved in the fission chain reaction of uranium-235. This task is performed by the control rods. A **control rod** contains material that can absorb neutrons. Cadmium and boron steel are commonly used in control rods. When a neutron strikes the nucleus of either of these atoms, it is absorbed into the nucleus and so takes no further part in the chain reaction.

Putting it all together

Figure 4.2.6(a) is a schematic diagram showing how all of the individual components of a nuclear reactor are combined. The **core** of a thermal nuclear reactor consists of the moderating material with fuel rods and control rods placed in it. These rods could be inserted into holes drilled into a pile of graphite several metres thick, or immersed into a volume of water or heavy water. The reactor that blew up at Chernobyl had about 1600 fuel rods and over 200 control rods in its graphite core.

When a neutron is released during the chain reaction, it is slowed by the moderating material. This enables it to be absorbed by a further uranium-235 nucleus, induce fission, and so continue the chain reaction. The rate of the chain reaction is controlled by raising or lowering the control rods. If the operators wish to reduce the energy output of the reactor, or even shut it down completely, they will lower the control rods further into the core. This has the effect of absorbing more neutrons and reducing or stopping the chain reaction.

The fission reaction in the reactor core produces an enormous amount of heat energy, resulting in the reactor core being typically at a temperature of 500–1500°C. This heat energy is removed from the core of the reactor by pipes that contain a **coolant** with a high specific heat capacity—typically liquid sodium, water, carbon dioxide gas, or heavy water. A **heat exchanger** then transfers this energy into pipes containing water. This water is converted into high-pressure steam that is used to rotate the turbines that drive the generator. The core of the reactor is encased in a protective **radiation shield** consisting of layers of concrete, steel, graphite and lead with a total thickness of about 2 metres. The function of this shield is to prevent neutrons and gamma rays from escaping from the reactor core, to protect the workers at the plant from damaging radiation.

PHYSICSFILE

Australia's nuclear reactor

Australia's only nuclear reactor is located at Lucas Heights, a suburb of Sydney. The HIFAR (High Flux Australian Reactor) operated here from 1958 to 2007 (Figure 4.2.4), but has now been replaced by a new research reactor known as the OPAL (Open Pool Australian Lightwater) reactor (Figure 4.2.5). About 7 kilograms of uranium enriched to 20% uranium-235 is immersed in a pool of water almost 13 metres deep. Radioisotopes that are used in industry and in the nuclear medicine departments of Australia's hospitals are synthesised here. Another important process that is carried out at Lucas Heights is the irradiation of silicon chips, which creates high-conductivity silicon for the computer industry.

FIGURE 4.2.4 The original housing of the HIFAR nuclear research reactor at Lucas Heights, Sydney decommissioned in 2007.

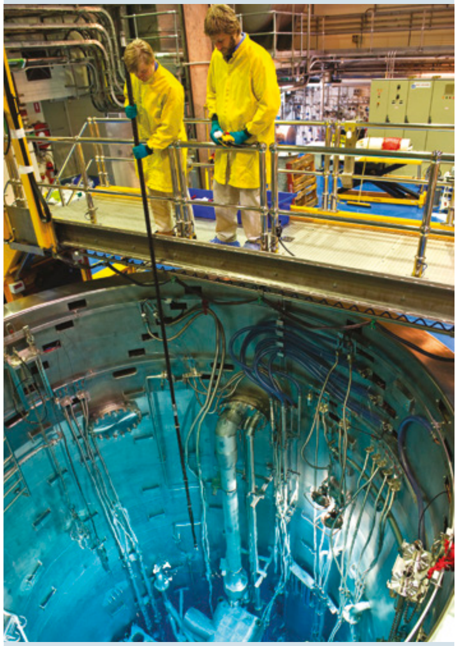
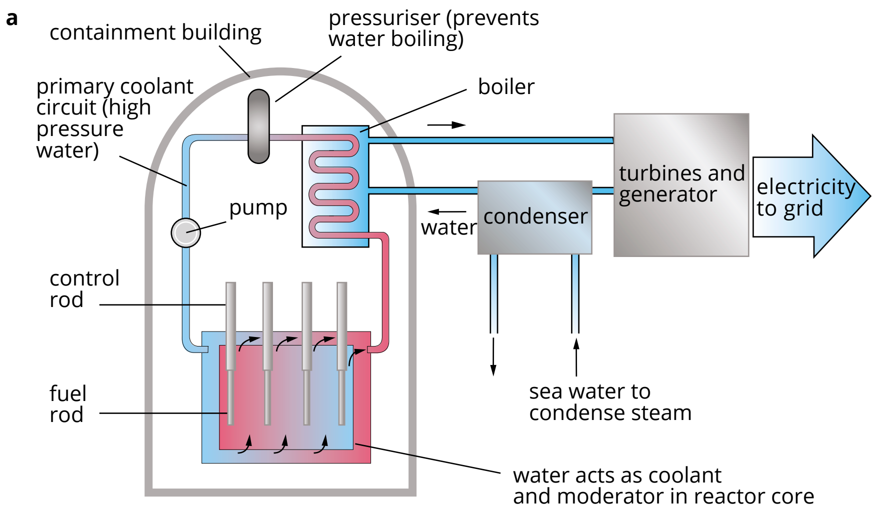


FIGURE 4.2.5 Inside the nuclear research reactor at Lucas Heights, Sydney.

The layers of graphite in the shield reflect escaping neutrons back into the core to take part in the chain reaction. The workers at a nuclear power plant are continually monitored to ensure that they are not exposed to unacceptably high levels of radiation, although their allowed dose is much higher than that of the general population.

Figure 4.2.6(b) shows the energy transformations that occur in a nuclear reactor in order to produce electricity. The primary difference between this and a gas- or coal-fired generator is in the way the heat is produced. A nuclear reactor uses the fission process and a gas or coal fired generator uses combustion of the fossil fuel. A typical 1000 MW power plant consumes about six million tonnes of black coal each year, or about 25 tonnes of enriched uranium that has been obtained from around 75 000 tonnes of ore.



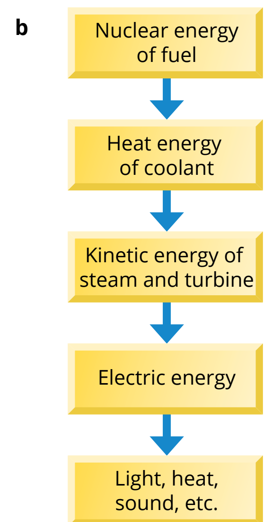


FIGURE 4.2.6 (a) A schematic diagram of a nuclear reactor; (b) energy transformations involved as a reactor is used to produce electricity.

PHYSICSFILE

Nuclear disasters: Fukushima

The biggest earthquake in Japan's history occurred on 11 March 2011 and triggered a massive tsunami along the east coast. Tragically, about 20 000 people were killed by the tsunami.

Japan has 52 nuclear reactors. The Fukushima nuclear power plant, consisting of six reactors, was close to the coast and protected by sea walls designed to withstand a 5.7 metre tsunami. The 14 metre wave that arrived about an hour after the earthquake completely inundated the reactors, electrical switching systems and generators. The pumps that pushed water through the reactors stopped working, causing the reactors to overheat. Hydrogen explosions caused fires and damaged the containment vessels. This resulted in the release of radioisotopes iodine-131 and caesium-137 into the atmosphere. Water leaked from Reactor 1, leaving fuel rods exposed, thus triggering their meltdown. As pictured below, one hydrogen explosion caused the roof of Fukushima's Reactor 1 to collapse, releasing radioactive materials into the atmosphere. The Japanese government created a 20 km exclusion zone around the power plant, forcing 80 000 people from their homes.

Nuclear accidents are rated according to the International Nuclear and Radiological Event scale from 1 to 7, with 7 being the most serious. Both the Fukushima and Chernobyl disasters (which occurred in Ukraine in 1986) were rated at 7.

FIGURE 4.2.7 A hydrogen explosion caused the roof of Reactor 1 at Fukushima to collapse, releasing radioactive materials into the atmosphere.

Management of nuclear waste

A major problem facing the nuclear power industry is the disposal of the unstable radioactive waste. There are about 400 nuclear power plants generating electricity around the world, producing large quantities of radioactive waste with long half-lives. A typical 1000 MW reactor will produce about 25 tonnes of spent fuel rods annually. The safe disposal of this high-level waste is of concern to many around the world.

Figure 4.2.8 shows the inside of the storage bunker for radioactive waste at Nieuwdorp in the Netherlands. All radioactive waste from across the Netherlands is collected and recorded then crushed into small packages and placed in 200-litre drums. This material is eventually decanted with concrete and placed in barrels for long-term storage. These barrels are then stored in a bunker secure against burglary and terrorist attacks.

FIGURE 4.2.8 The radioactive waste storage bunker in Nieuwdorp, Netherlands.

Radioactive waste products are classified as low-level, intermediate level or high-level waste.

- Low-level waste is generated primarily from hospitals, industry and laboratories and consists mostly of tools, clothing, used wrapping material and other items that have been contaminated with radionuclides with short half-lives. Low-level waste solids are usually compacted or incinerated, then buried in shallow pits on land or at sea.
- Intermediate-level waste typically consists of reactor components, chemical sludges, and contaminated materials from reactors that have been decommissioned. Intermediate-level wastes are solidified in bitumen or concrete, then buried or stored in deep trenches.
- High-level waste is waste from contaminated reactor parts, as well as liquid waste from the reprocessing of spent fuel rods. This waste contains highly radioactive fission fragments and transuranic elements, and so requires special shielding during handling and transport. As can be seen in the graph in Figure 4.2.9, high-level waste remains radioactive for an exceedingly long time and needs to be stored permanently. By way of comparison, the activity of one tonne of uranium ore is only 8×10^{11} Bq.

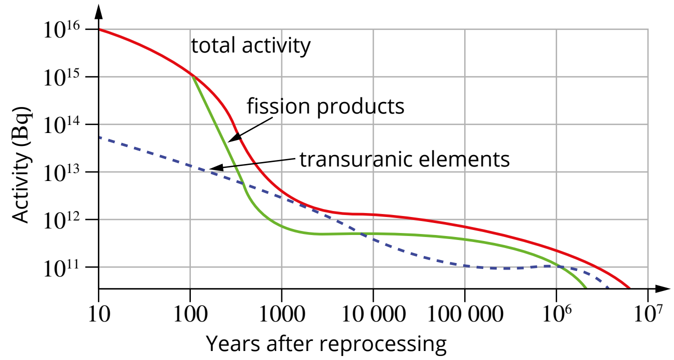


FIGURE 4.2.9 The activity of 1 tonne of high-level nuclear waste.

In the USA and Canada, spent fuel rods are permanently stored in cooling ponds. The uranium-235 and uranium-238 nuclei in these fuel rods have half-lives of 700 000 and 4.5 billion years respectively. The numerous fission fragments have a wide range of half-lives. The fuel rods must be permanently stored in cooling ponds inside nuclear power stations in order to protect people from the radioactive emissions. In Japan, Russia and Europe, the spent fuel rods are reprocessed. The uranium is extracted and reused as nuclear fuel. This is an expensive process.

4.2 Review

SUMMARY

- A nuclear reactor uses fuel rods that contain uranium-235 enriched to between 3% and 4% as its fuel.
- The core of the reactor contains a material called a moderator, such as graphite or water. This material slows neutrons down, enabling them to be captured by the uranium-235 nucleus and bring about a fission reaction.
- The rate at which the fission reactions occur is determined by control rods. These are raised and lowered in the core and contain a material that absorbs neutrons, such as cadmium or boron, and thereby controls the chain reaction.
- The coolant is a liquid that flows through the core of the reactor extracting heat from it and transferring this, through a heat exchanger, to water that is converted to steam to drive the generator turbines.
- Uranium-238 absorbs neutrons and undergoes transmutation to produce plutonium-239 as one of its daughter nuclei.
- Plutonium-239 is used as a fuel in fast breeder reactors as it undergoes fission when struck with fast-moving neutrons.
- A chain reaction may be established when a fission reaction occurs that generates one or more neutrons per fission
- Critical mass is the minimum amount of enriched fissile material in the shape of a sphere required to sustain a chain reaction.
- Many of the fission products from a nuclear reactor have very long half-lives. This makes the long term safe storage of this waste difficult.

KEY QUESTIONS

1 Which one of the following isotopes is most suitable as a fuel for a nuclear reactor?
   A U-238
   B U-235
   C U-234
   D Pu-239

2 The uranium ore that is dug from the ground contains two different isotopes, uranium-235 and uranium-238. Explain why uranium in this form is not immediately suitable as a fuel for a nuclear reactor.

3 The uranium that is used as the fuel for a nuclear reactor has been enriched so that its uranium-235 content is around:
   A 0.7%
   B 4%
   C 10%
   D 95%

4 Describe the function of the moderator in a nuclear reactor.

5 Describe the function of the control rods in a nuclear reactor.

6 In a chain reaction that results in an explosion, approximately 10^{24} uranium-235 nuclei undergo fission in just over 1 μs. What conditions would be necessary for this to occur?

7 The critical mass of uranium-235 is about 1 kg, but a 5 kg piece of uranium-235 that is flattened like a sheet is not capable of exploding. Explain why.

8 Explain why lead ($Z = 82$) would be unsuitable for use as a moderator.

9 Describe the effect on the operation of a nuclear reactor if the number of neutrons released per fission is:
   a equal to one
   b less than one
   c greater than one.

10 The fissile material that is used in a nuclear reactor is uranium-235. What is the effect on the nuclei of this isotope when it is bombarded with:
   a fast neutrons?
   b slow neutrons?

11 Approximately 97% of the uranium in the fuel rods of nuclear reactors is uranium-238. When struck by a fast neutron, a uranium-238 nucleus is more likely than a uranium-235 nucleus to absorb the neutron.
   a In what way does this change the uranium nuclei?
   b Why does this lead to problems in disposing of the nuclear waste from the reactor?

12 During the fission of plutonium-239, the average number of neutrons released is 2.91 per fission. This is higher than the average of 2.47 released during the fission of uranium-235. How is this of benefit in a fast breeder reactor?

13 The neutron bombardment of uranium-238 triggers two successive beta decays before reaching the final product of plutonium-239. The fission of plutonium-239 randomly produces many pairs of nuclei with one of these being xenon-134 and zircon-103.
   a Write equations for the three steps between U-238 and Pu-239.
   b Write an equation for the fission of Pu-239 to Xe-134 and Zr-103.

14 During the lifetime of a reactor, the control rods need to be gradually removed over a period of months to maintain the energy production at a constant rate. Explain why this is necessary.

4.3 Nuclear fusion

Nuclear fusion is a process that has been occurring inside the Sun and other stars for billions of years. **Fusion** involves the combining of small nuclei such as hydrogen and helium to form a larger nucleus. The amount of energy released per nucleon with fusion is greater than with fission and there is no radioactive waste produced. Scientists are working on experimental fusion reactors such as ITER (International Thermonuclear Experimental Reactor) in France and the National Ignition Facility in the USA, which is shown in Figure 4.3.1.

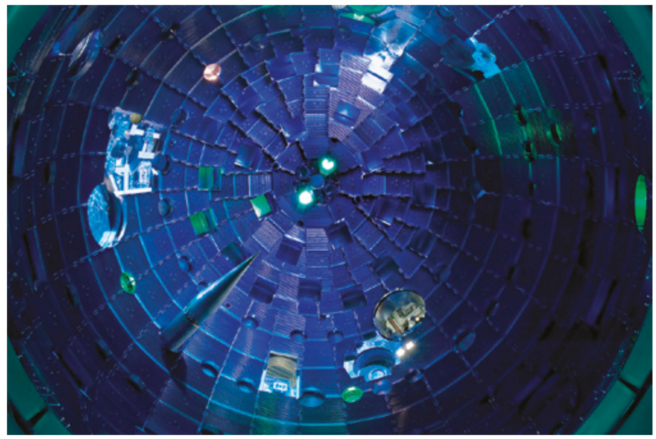


FIGURE 4.3.1 The experimental fusion reactor at the National Ignition Facility in the USA.

The fusion process created so far has only lasted for a fraction of a second and it is not expected that a fusion reactor will successfully operate for many years. Researchers at Lockheed Martin in the USA are working on a compact fusion reactor. In 2014 they claimed a prototype will be running by 2019. This claim has been met with scepticism by some in the scientific community.

In the study of nuclear reactions, Einstein's idea of mass–energy equivalence is used to explain why energy is released during the splitting of a nucleus. When fission occurs, the mass of the particles in the fission fragments is lower than that of the parent nucleus. Einstein's famous equation, $E = mc^2$, can be used to calculate the amount of energy that is related to this missing mass. This equation also applies to fusion reactions. It is important to note that less than 1% of matter is converted into energy in all of the energy transformations being discussed here.

NUCLEAR FUSION

Nuclear fusion occurs when two light nuclei are combined (fuse) to form a larger nucleus. The example of nuclei fusing to form a helium atom is shown in Figure 4.3.2(a).

As in the cases of radioactive decay and nuclear fission, the mass of the reactants is slightly greater than the mass of the products when the nuclei combine during fusion. This mass difference is represented by the unbalanced scales shown in Figure 4.3.2(b).

The energy created by this missing mass can again be determined from:

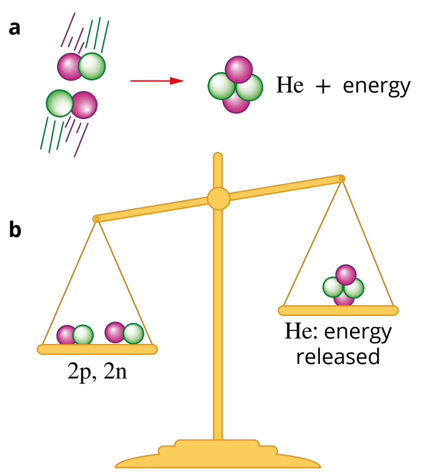


FIGURE 4.3.2 (a) When two isotopes of hydrogen fuse to form a helium nucleus, energy is released. (b) The loss in mass, Δm, can be calculated using $\Delta E = \Delta mc^2$.

$\Delta E = \Delta mc^2$

where ΔE is energy (J)

Δm is the mass defect (the decrease in mass, in kg)

c is the speed of light $= 3.0 \times 10^8$ m s^{-1}

$\Delta E = \Delta u \times 931$

or

where ΔE is energy (MeV)

Δu is the mass defect (in atomic mass units)

931 is the mass-energy equivalent

Nuclear fusion is very difficult to achieve. The main problem is that nuclei are positively charged. They exert an electrostatic force of repulsion on each other—that is, they push each other away. Therefore it is not easy to force the nuclei together. Remember that the electrostatic force is a long-range force and the strong nuclear force of attraction only acts at much shorter distances.

As two nuclei approach each other, the electrostatic force will cause them to be repelled. Slow-moving nuclei with relatively small amounts of kinetic energy will not be able to get close enough for the strong nuclear force to come into effect. Fusion will not happen, as can be seen in Figure 4.3.3(a).

If the nuclei travel towards each other at higher speeds, as shown in Figure 4.3.3(b), they may have enough kinetic energy to overcome the repulsive force. The nuclei can now get close enough for the strong nuclear force to start acting. If this happens, fusion will occur.

The graph in Figure 4.3.4 shows the effect of the electrostatic force and the strong nuclear force on the potential energy of a pair of deuterium ($^{2}_{1}H$) nuclei. At large separation distances, the electrostatic force dominates and the nuclei repel each other (shown to the right of the energy barrier in the graph). At small separation distances, the strong nuclear force dominates and the nuclei can fuse together. However, to get the nuclei to this point, they need an enormous amount of energy. Temperatures in the hundreds of millions of degrees are required (as in the Sun). This enormous amount of energy enables the nuclei to overcome the energy barrier shown in the graph and fuse together.

As in fission, in any fusion reaction the atomic numbers and mass numbers on either side of the equation are conserved. The fusion of two hydrogen-2 (deuterium) nuclei is shown below:

$$^{2}_{1}H + ^{2}_{1}H \rightarrow ^{3}_{1}H + ^{1}_{1}H$$

The atomic numbers add up to two on both sides and the mass numbers add up to four on both sides. However, the total mass of the reactants (on the left-hand side) will be greater than the total mass of the products (on the right-hand side). The difference in mass—the mass defect—can be measured in either kilograms (kg) or atomic mass units (u), where $1\,u = 1.6606 \times 10^{-27}$ kg. If mass it is expressed in kg then the energy in joules is calculated using $\Delta E = \Delta mc^2$. If the mass is expressed in atomic mass units then the energy can be calculated directly using the mass-energy equivalent conversion that $1\,u = 931$ MeV.

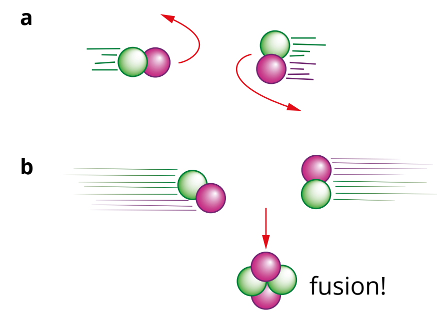


FIGURE 4.3.3 (a) Slow-moving nuclei do not have enough energy to fuse together. The electrostatic forces cause them to be repelled from each other. (b) If the nuclei have sufficient kinetic energy, then they will overcome the repulsive forces and move close enough together for the strong nuclear force to come into effect. At this point, fusion will occur and energy will be released.

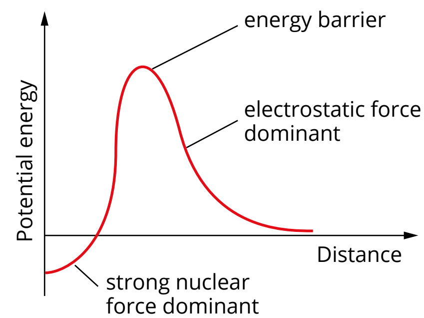


FIGURE 4.3.4 If two hydrogen-2 (deuterium) nuclei are to get close enough for the strong nuclear force to act, they must overcome the energy barrier presented by the electrostatic force.

PHYSICSFILE

Hydrogen bomb

In 1952, a fusion reaction was used to power the world's first hydrogen bomb. It had five times the destructive power of all the conventional bombs that were dropped during the whole of World War II. The high temperature achieved by a fissile fuel explosion was used to initiate the fusion reaction. In other words, an atomic bomb was used as the fuse for the hydrogen fusion bomb.

FIGURE 4.3.5 The hydrogen bomb dropped at Bikini Atoll in 1956.

Worked example 4.3.1

FUSION

Two deuterium nuclei undergo fusion to produce one tritium nucleus, a proton and energy according to the equation below. Calculate the energy released in this reaction. Use the following data in your calculations: mass of deuterium = 2.013553 u, mass of tritium = 3.016049 u and mass of a proton is 1.007276 u. ${}^{2}_{1}H + {}^{2}_{1}H \rightarrow {}^{3}_{1}H + {}^{1}_{1}H + \text{energy}$	
Thinking	**Working**
Determine the mass of the reactants.	mass of reactants = 2 × mass of deuterium = 2 × 2.013553 = 4.027106 u
Determine the mass of the products.	mass of products = mass of tritium + mass of proton = 3.016049 + 1.007276 = 4.023325 u
Determine the mass defect.	mass defect = mass of reactants – mass of products = 4.027106 – 4.023325 = 0.0037807 u
Determine the energy equivalent.	= mass defect × 931 MeV = 0.0037807 × 931 = 3.521 MeV
Convert to joules.	$= 3.521 \times 10^{6} \times 1.60 \times 10^{-19}$ $= 5.63 \times 10^{-13}$ J

Worked example: Try yourself 4.3.1

FUSION

One of the possible nuclear fusion reactions in a star involves the fusion of two helium-3 nuclei to produce a helium-4 nucleus, two protons and energy according to the equation below. Calculate the energy, in joules and MeV, released in this reaction. Use the following data in your calculations: mass of helium-3 nucleus = 3.014932 u, mass of helium-4 nucleus = 4.001505 u and mass of a proton = 1.007276 u. ${}^{3}_{2}He + {}^{3}_{2}He \rightarrow {}^{4}_{2}He + 2{}^{1}_{1}H + \text{energy}$

BINDING ENERGY

As seen earlier in this chapter, the total mass of a stable nucleus is slightly less than the combined mass of the individual protons and neutrons. Einstein realised that this mass defect, Δm, is converted to energy using: $\Delta E = \Delta mc^2$. This is known as the binding energy of the nucleus. The binding energy of the nucleus indicates how much energy is needed to separate the nucleus into individual protons and neutrons.

Each nucleus has its own binding energy value, with a higher value indicating a more stable nucleus. A binding-energy-per-nucleon graph, as shown in Figure 4.3.6, allows a comparison of nuclear stabilities.

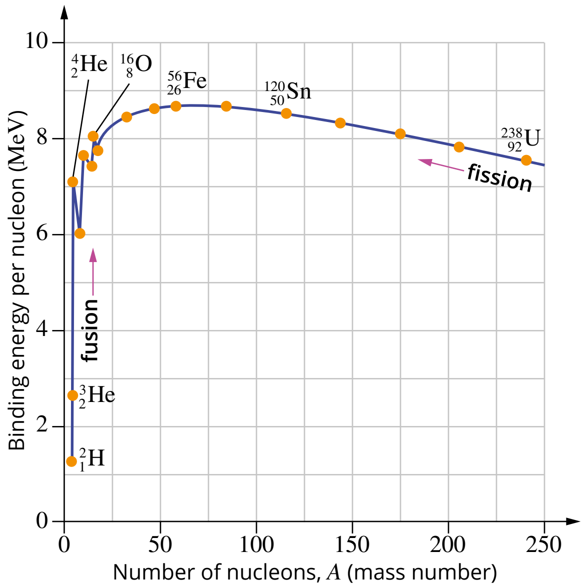


FIGURE 4.3.6 The graph of binding energy per nucleon.

Figure 4.3.6 can be analysed to better understand fission and fusion.

- Small nuclei have lower binding energy per nucleon values, indicating that they are easier to break apart compared to larger nuclei. Helium-4 has a relatively high value indicating that it is stable.
- The binding energy per nucleon increases dramatically for very small nuclei. As they fuse together, the binding energy per nucleon increases. This is the energy released during fusion.
- Elements with mass numbers between 40 and 80 have nuclei that are tightly bound. It takes more energy to break these nuclei apart. These are the most stable nuclei. As you can see from the graph, these elements have the highest binding energy per nucleon. The higher the binding energy per nucleon, the more stable the nucleus.
- Larger nuclei have lower binding energy per nucleon values, indicating that they are less stable.
- If a large nucleus such as uranium splits into two fragments, the binding energy per nucleon of the fragments again increases. This is the energy released during fission.
- Iron (Fe), with a mass number of 56, has the most stable nucleus.
- Nuclei smaller than iron undergo fusion and release energy. Nuclei larger than iron undergo fission and release energy.

Worked example 4.3.2

BINDING ENERGY

Calculate the average binding energy per nucleon for the carbon-12 nucleus in MeV and joules. Use the following data in your calculations: mass of a carbon-12 nucleus = 11.993417 u, mass of a proton = 1.007276 u and mass of a neutron = 1.008664 u.

Thinking	Working
Determine the total mass of the nucleons in a carbon-12 nucleus.	total mass = mass of 6 neutrons + mass of 6 protons $= (6 \times 1.008664) + (6 \times 1.007276)$ $= 12.095640\ \text{u}$
Determine the mass defect.	= mass of nucleons – actual mass of nucleus $= 12.09564 - 11.993417$ $= 0.102223\ \text{u}$
Determine the binding energy in MeV.	= mass defect × 931 MeV $= 0.102223 \times 931$ $= 95.17\ \text{MeV}$
Determine the binding energy per nucleon in MeV.	$= \frac{95.17}{12}$ = 7.93 MeV per nucleon
Determine the binding energy per nucleon in J.	$= 7.93 \times 10^{6} \times 1.60 \times 10^{-19}$ $= 1.27 \times 10^{-12}\ \text{J}$

Worked example: Try yourself 4.3.2

BINDING ENERGY

Calculate the average binding energy per nucleon for the uranium-235 nucleus in MeV and joules. Use the following data in your calculations: mass of a uranium-235 nucleus = 234.993462 u, mass of a proton = 1.007276 u and mass of a neutron = 1.008664 u.

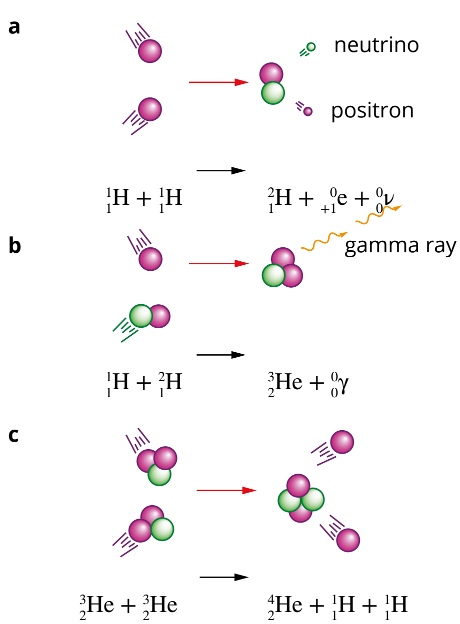


FIGURE 4.3.7 The three main fusion reactions taking place inside in the Sun.

FUSION IN THE SUN AND SIMILAR STARS

In the Sun, many different fusion reactions are taking place. The main reaction is the fusion of hydrogen nuclei to form helium. Each second, about 657 million tonnes of hydrogen and hydrogen isotopes fuse to form about 653 million tonnes of helium. Each second, a mass defect of 4 million tonnes results from these fusion reactions. The amount of energy released is enormous and can be found by using the equation $\Delta E = \Delta mc^2$. A tiny proportion of this energy reaches Earth and sustains life.

The sequence of fusion reactions shown in Figure 4.3.7 has been occurring inside the Sun for the past 5 billion years and is expected to last for another 5 billion years or so. Hydrogen nuclei are fused together and, after several steps, a helium nucleus is formed. This process releases about 25 MeV of energy.

NUCLEAR FUSION REACTORS

You might recall from studying radioactivity that for each nuclear fission reaction, about 200 MeV of energy is released. This is a lot of energy relative to other nuclear reactions, but fission typically involves nuclei with around 240 nucleons.

In nuclear fusion, less energy is emitted—around 20–25 MeV. However, this energy is released from a reaction involving just a few nucleons. The energy per nucleon is greater than that for fission reactions because a greater percentage of the mass is transformed into energy. For scientists working on experimental fusion reactors, this is very important as massive amounts of energy could be obtained from fusion reactions between small, readily available nuclides with no radioactive by-products.

Since the 1950s, a great deal of research has been devoted to recreating nuclear fusion in a laboratory. However, major technical problems have been encountered in trying to initiate ongoing fusion reactions on Earth. Replicating the reactions that take place on the Sun is extraordinarily difficult. This is because extremely high densities, temperatures and pressures are required. Fusion researchers currently use two isotopes of hydrogen—deuterium and tritium—as fuel. Deuterium can be extracted in vast quantities from lakes and oceans, but tritium is radioactive, with a half-life of 12.3 years, and must be artificially produced. The nuclear reactions used in current fusion reactors are as follows:

$$^{2}_{1}\mathrm{H} + ^{2}_{1}\mathrm{H} \rightarrow ^{3}_{2}\mathrm{He} + ^{1}_{0}\mathrm{n}$$

$$^{2}_{1}\mathrm{H} + ^{2}_{1}\mathrm{H} \rightarrow ^{3}_{1}\mathrm{H} + ^{1}_{1}\mathrm{H}$$

$$^{2}_{1}\mathrm{H} + ^{3}_{1}\mathrm{H} \rightarrow ^{4}_{2}\mathrm{He} + ^{1}_{0}\mathrm{n}$$

The main obstacle to the development of a successful fusion reactor is the extremely high temperature that must be achieved before fusion can commence. Temperatures of over 100 million degrees are needed to trigger a self-sustaining fusion reaction and this must then be contained inside the reactor. Current fusion reactors have achieved temperatures of about 100 million degrees, but only for very short periods.

PHYSICSFILE

Tunnel effect

The temperature at the surface of the Sun is about 5500°C, hot enough to vaporise any known material. The Sun is much, much hotter at the core, where temperatures reach around 15 million degrees. But this is well below the hundreds of millions of degrees needed to sustain nuclear fusion. The fact that fusion occurs at these low temperatures cannot be explained by classical physics. Quantum mechanics must be used—treating the protons as waves rather than particles. This process is called the tunnel effect and is beyond the scope of this course.

EXTENSION

Fusion-reactor research

A commercial nuclear fusion reactor is perhaps four or five decades into the future. At the Joint European Torus, or JET, in England, research is being carried out into tokamaks—doughnut-shaped reactors that use magnetic fields to contain the plasma of fusion reactants away from the reactor walls and can reach temperatures of hundreds of millions of degrees in the torus. The design of the reactor and shape of the torus are shown in figures 4.3.8 and 4.39. The largest project in this area is being conducted at the ITER (International Thermonuclear Experimental Reactor). This experiment is a joint venture between Europe, China, India, Japan, Russia, the USA and South Korea, and it is being constructed in France. The aim is to initiate plasma experiments by 2020.

FIGURE 4.3.9 The doughnut shape or torus of the JET fusion reactor can be clearly seen in this photograph.

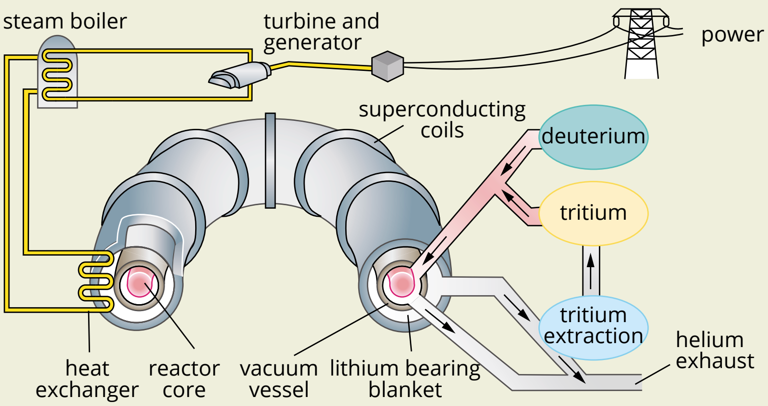


FIGURE 4.3.8 Design of experimental tokamak reactor.

Another technique called inertial confinement is being investigated at the National Ignition Facility in California, USA. Here, a pellet of fusion fuel is zapped by powerful lasers, causing the fuel to implode. This implosion initiates the fusion reaction.

Nuclear fusion may perhaps be the ultimate energy solution. The fuel for nuclear fusion, deuterium, can be readily obtained from seawater. Vast quantities of energy are released during the fusion process, yet a relatively small amount of radioactive waste is created. The radioactive waste consists mostly of the reactor parts that have suffered neutron irradiation.

4.3 Review

SUMMARY

- Nuclear fusion is the combining of light nuclei to form heavier nuclei. Extremely high temperatures are required for fusion to occur.
- Nuclear fusion is the process producing energy in the Sun and other stars.
- Approaching nuclei must have enough energy to overcome the electrostatic forces and get close enough for the strong nuclear force to take effect. The energy that is required to achieve this is called the energy barrier.
- The mass of the product nuclei is less than the combined mass of the parent nuclei, and this mass difference accounts for the energy released in the reaction according to Einstein's equation $\Delta E = \Delta mc^2$.
- The amount of energy released per nucleon is greater for fusion than it is for fission.
- The binding energy indicates how much energy is needed to separate the nucleus into individual protons and neutrons.
- If very small nuclei fuse together, their binding energy per nucleon increases. For very small nuclei, this increases dramatically.
- If a large nucleus such as uranium splits into two fragments, the binding energy per nucleon in the fragments increases.
- Very high temperatures are needed to overcome the repulsive forces between the nuclei in order to trigger and sustain a nuclear reaction. This is the main obstacle to the development of fusion reactors.
- Fusion creates very little radioactive waste.

KEY QUESTIONS

1 How does fusion differ from fission?

2 The fusion reaction that is most promising for use in nuclear fusion reactors is: ${}^{2}_{1}\text{H} + {}^{3}_{1}\text{H} \rightarrow {}^{4}_{2}\text{He} + {}^{1}_{0}\text{n}$.
Why is energy released during this process?

3 How does the amount of energy released per nucleon during a single nuclear fission reaction compare to the amount of energy released per nucleon for a single fusion reaction?

4 During the process of nuclear fusion, mass is lost and this appears as energy. What is the approximate percentage of mass lost during a typical fusion reaction?

5 Consider the fusion reaction shown below.
${}^{2}_{1}\text{H} + {}^{3}_{1}\text{H} \rightarrow {}^{a}_{b}X + {}^{1}_{0}\text{n}$

- **a** Determine the values of *a* and *b* and hence the symbol of the unknown element X.
- **b** 33 MeV of energy is released. What is the mass defect in kg? Give your answer to two significant figures and in scientific notation.

6 Two slow-moving protons are travelling directly towards each other. Explain whether or not the protons will collide and fuse together.

7 Two fast-moving protons are travelling directly towards each other. The protons collide and fuse together. Explain why this happens.

8 The following fusion reaction is taking place in the Sun.
${}^{3}_{2}\text{He} + X \rightarrow {}^{4}_{2}\text{He} + {}^{1}_{1}\text{p} + {}^{1}_{1}\text{p}$
During each fusion reaction, 23 MeV of energy is released. Give your answers to two significant figures where appropriate and in scientific notation.

- **a** What are the atomic and mass numbers of particle *X* and what is its symbol?
- **b** Convert 23 MeV into joules.
- **c** Determine the mass defect for this fusion process.

9 What happens to the binding energy and the stability of two hydrogen-2 nuclei when they are fused together to form helium-4?

10 What happens to the number of nucleons during the fusion reaction below?
${}^{2}_{1}\text{H} + {}^{3}_{1}\text{H} \rightarrow {}^{4}_{2}\text{He} + {}^{1}_{0}\text{p}$

Chapter review

04

KEY TERMS

binding energy	electronvolt	heat exchanger
chain reaction	fast breeder reactor	heavy water
control rod	fissile	moderator
coolant	fission	radiation shield
core	fission fragments	subcritical mass
critical mass	fuel rod	supercritical mass
daughter nuclei	fusion	transuranic

1 What does the term 'fissile' mean?

2 Are all atoms fissile? Give examples to support your answer.

3 Consider one particular proton in the nucleus of a zinc atom. Describe the forces that the proton experiences from other nucleons.

4 Neutrons and alpha-particles can both be used to trigger nuclear fission. Explain why neutrons are better than alpha-particles for inducing fission.

5 Einstein said that mass and energy are equivalent. In one particular fission reaction, a decrease in mass of 3.48×10^{-28} kg occurs. Give your answers to two significant figures.

a Express the energy equivalent of this mass in joules.

b Express the energy equivalent of this mass in MeV.

6 Determine the value of the unknown mass number *x* and atomic number *y* in this fission reaction:

$^1_0n + ^x_{94}Pu \rightarrow ^{130}_{54}Xe + ^{106}_{y}Xr + 4^1_0n$

7 Determine the number of neutrons (*x*) released in the following fission reaction:

$^1_0n + ^{235}_{92}U \rightarrow ^{127}_{50}Sn + ^{102}_{42}Mo + x^1_0n$

8 A typical fusion reaction is: $^2_1H + ^2_1H \rightarrow ^3_1H + ^1_1H$

Why are high temperatures such as 100 million degrees needed for this reaction to occur?

9 Consider the following fission reaction of uranium-235:

$^1_0n + ^{235}_{92}U \rightarrow ^{141}_{55}Cs + ^{93}_{37}Rb + 2^1_0n$

During this reaction there is a mass defect of 4.99×10^{-28} kg. How much energy in joules is produced per reaction?

10 Consider the following fusion reaction:

$^1_1H + ^3_2He \rightarrow ^4_2He + ^0_{+1}e + \nu$

Hydrogen and helium-3 are fused together and a helium-4 nucleus is created along with a positron and a neutrino. 21 MeV of energy is released.

a How does the combined mass of the hydrogen and helium-3 nucleus compare with the combined mass of the helium-4 nucleus, positron and neutrino?

b Where has the energy that was released come from?

c Convert the energy into joules.

d What is the mass defect of this fusion reaction?

11 Compare the waste and the energy per nucleon produced by fusion reactors with those of fission reactors.

12 What happens to the binding energy per nucleon and the stability of the nucleus when a uranium-238 nucleus splits apart to form two smaller nuclei?

13 The binding energy per nucleon for iron (mass number 56) is higher than for other elements. What does this mean for the stability of iron nuclei?

14 Which form of electromagnetic radiation is released from the nucleus of radioactive atoms?

15 Plutonium is an element that does not exist naturally on Earth. How can this element be available for use in nuclear reactors?

16 Explain the primary function of each of the following components of a nuclear reactor, and state what material they are made of:

a the coolant **b** the moderator

c the control rods.

17 In a tokomak fusion reactor a proton fuses with a deuterium nucleus to form a helium nucleus and release a gamma ray. The energy created during this process is 20.0 MeV.

a Write the equation for this reaction.

b How much energy is released per reaction in joules?

c Calculate the mass defect for this reaction.

18 A series of fusion reactions that occur in the Sun begins with two protons fusing to form deuterium and a positron. A deuterium nucleus and another proton can then fuse to form a helium-3 nucleus and a gamma ray. Two helium-3 nuclei can fuse to form an alpha particle and two protons with the release of 12.98 MeV of energy per alpha particle formed.

a Write equations for the three fusion reactions stated above.

b If the last reaction of these three could be controlled in a reactor, how much power could be produced from the fusion, in one day, of 100 g of helium-3 into alpha particles and protons? Use the following masses in your calculations: helium-3 nucleus = 3.01493 u, helium-4 nucleus = 4.00151 and proton = 1.00728, where 1 u = 1.66054×10^{-27} kg.

CHAPTER

05 Electrical physics

Every object around you is made up of charged particles. When these particles move relative to one another, we experience a phenomenon known as 'electricity'. This chapter looks at the fundamental concepts such as current and potential difference that scientists have developed to explain electrical phenomena. This will provide the foundation for studying practical electrical circuits in the following chapter.

Science as a Human Endeavour

The supply of electricity to homes has had an enormous impact on society and the environment. An understanding of electrical circuits informs the design of effective safety devices for the safe operation of: lighting, power points, stoves and other household devices

Science Understanding

- there are two types of charge that exert forces on each other
- electric current is carried by discrete charge carriers; charge is conserved at all points in an electrical circuit. *This includes applying the relationship* $I = \frac{q}{t}$
- energy is conserved in the energy transfers and transformations that occur in an electrical circuit
- the energy available to charges moving in an electrical circuit is measured using electric potential difference, which is defined as the change in potential energy per unit charge between two defined points in the circuit. *This includes applying the relationship* $V = \frac{W}{q}$
- energy is required to separate positive and negative charge carriers; charge separation produces an electrical potential difference that drives current in circuits
- power is the rate at which energy is transformed by a circuit component; power enables quantitative analysis of energy transformations in the circuit. *This includes applying the relationship* $P = \frac{W}{t} = VI$
- resistance depends upon the nature and dimensions of a conductor
- resistance for ohmic and non-ohmic components is defined as the ratio of potential difference across the component to the current in the component. *This includes applying the relationship* $R = \frac{V}{I}$
- circuit analysis and design involve calculation of the potential difference across the current in, and the power supplied to, components in series, parallel, and series/parallel circuits. *This includes applying the relationships*

  series components, I = constant, $V_t = V_1 + V_2 + V_3 \ldots$

  $R_t = R_1 + R_2 + R_3 \ldots$

  parallel components, V = constant, $I_t = I_1 + I_2 + I_3 \ldots$

  $$\frac{1}{R_t} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \ldots$$
- there is an inherent danger involved with the use of electricity that can be reduced by using various safety devices, including fuses, residual current devices (RCD), circuit breakers, earth wires and double insulation
- electrical circuits enable electrical energy to be transferred and transformed into a range of other useful forms of energy, including thermal and kinetic energy, and light

5.1 Behaviour of charged particles

All matter in the universe is made of tiny particles. These particles have a property called **charge** that can be either positive, negative, or neutral. Usually, the numbers of positive and negative charges balance out so perfectly that we are completely unaware of them. However, when significant numbers of these charged particles become separated or move relative to each other, it results in **electricity**.

In order to understand electricity, it is important to first understand the way charged particles interact with each other.

EXISTENCE OF CHARGE CARRIERS

The tiny particles that make up all matter are called atoms. Every atom contains a nucleus at its centre, made up of positively charged particles called protons and neutral particles called neutrons. The nucleus, which is positively charged due to the protons, is surrounded by negatively charged electrons. A model of an atom is shown in Figure 5.1.1.

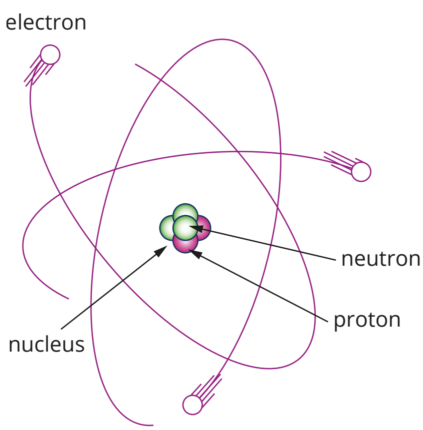


FIGURE 5.1.1 A simple model of an atom.

Simple models of the atom, often called planetary models, show the electrons as orbiting the nucleus much like the planets orbit the Sun. This is because particles that have the same charge will repel each other, while particles with opposite charges will attract each other. In an atom the negatively charged electrons are held in place by their attraction to the positively charged nucleus.

This is an important rule to remember when thinking about the interaction of charged particles.

Charge	Positive	Negative
positive	repel	attract
negative	attract	repel

In neutral atoms, the number of electrons is exactly the same as the number of protons. This means that their charges balance each other out, leaving the atom electrically neutral.

It is difficult to remove a proton from the nucleus of an atom. In comparison, the outer electrons of an atom are relatively loosely held and so they are relatively easy to remove.

PHYSICSFILE

Electron models

The way in which an electron moves around the nucleus of an atom is more complex than the simple planetary model would suggest. An individual electron is so small that its exact position at any point in time is impossible to measure. Recent models of the structure of the atom describe an electron in terms of the probability of finding it in a certain location. In diagrams of atoms, this is often represented by drawing the electrons around the nucleus as a fuzzy cloud, rather than points or solid spheres.

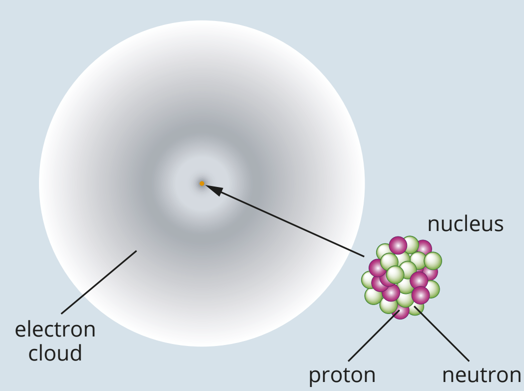


FIGURE 5.1.2 The nucleus of an atom occupies about 10^{-12} of the volume of the atom, yet it contains more than 99% of its mass. Atoms are mostly empty space.

When electrons move from one object to another, each object is said to have gained a **net charge**. The object that loses the electrons will have a net positive charge, since it will now have less negative electrons than positive protons. The object that gains electrons will have a net negative charge. When an atom has gained or lost electrons, you say it has been **ionised** or has become an **ion**.

The understanding that the movement of electrons, rather than protons, creates electrical effects is a relatively new discovery. Unfortunately, this means that many of the rules and conventions used when talking about electricity refer to electric current as the movement of positive charge carriers.

> An excess of electrons causes an object to be negatively charged, and a deficit in electrons will mean the object is positively charged.

MEASURING CHARGE

In order to measure the actual amount of charge on a charged object, a 'fundamental' unit would be the size of the charge on one electron or one proton. This fundamental charge is often referred to as the **elementary charge** and is given the symbol e. A proton therefore has a charge of $+e$ and an electron has a charge of $-e$.

The Systeme International (SI) unit of charge is known as the **coulomb** (symbol C). It is named after Charles-Augustin de Coulomb, who was the first scientist to measure the forces of attraction and repulsion between charges.

A coulomb is quite a large unit of charge: +1 coulomb (1 C) is equivalent to the combined charge of 6.2×10^{18} protons. The size of the elementary charge is very small, the charge on a single proton is $+1.6 \times 10^{-19}$ C and the charge on a single electron is -1.6×10^{-19} C. The letter q is used to represent the quantity of charge.

> The elementary charge, e, of a proton is equal to $+1.6 \times 10^{-19}$ C.
>
> The elementary charge, $-e$, of an electron is equal to -1.6×10^{-19} C.

Worked example 5.1.1

THE AMOUNT OF CHARGE ON A GROUP OF ELECTRONS

Calculate the charge, in coulombs, carried by 6 billion electrons.

Thinking	Working
Express 6 billion in scientific notation.	1 billion = 10^9 6 billion = 6×10^9
Calculate the charge, Q, in coulombs by multiplying the number of electrons by the charge on an electron (-1.6×10^{-19} C).	$q = (6 \times 10^9) \times (-e)$ $= (6 \times 10^9) \times (-1.6 \times 10^{-19}\,C)$ $= -9.6 \times 10^{-10}\,C$

Worked example: Try yourself 5.1.1

THE AMOUNT OF CHARGE ON A GROUP OF ELECTRONS

Calculate the charge, in coulombs, carried by 4 million electrons.

Worked example 5.1.2

THE NUMBER OF ELECTRONS IN A GIVEN AMOUNT OF CHARGE

The net charge on an object is $-3.0\,\mu C$ ($1\,\mu C$ = 1 microcoulomb = 10^{-6} C). Calculate the number of extra electrons on the object.

Thinking	Working
Express $-3.0\,\mu C$ in scientific notation.	$q = -3.0\,\mu C$ $= -3.0 \times 10^{-6}\,C$
Find the number of electrons by dividing the charge on the object by the charge on an electron (-1.6×10^{-19} C).	$n_e = \frac{q}{-e}$ $= \frac{-3.0 \times 10^{-6}\,C}{-1.6 \times 10^{-19}\,C}$ $= 1.9 \times 10^{13}$ electrons

PHYSICSFILE

Separating positive and negative charges

Electrons can be **transferred** (moved) from one object to another by simply rubbing two objects together. The objects need to be made from different materials. You can see this if you rub a balloon against your hair and then slowly move the balloon away. You will notice that your hair seems to stick to the balloon. This is because electrons are rubbed off your hair and transferred onto the balloon. This causes the balloon to gain a net negative charge and your hair to gain a net positive charge, which means the balloon and your hair are attracted to each other.

Worked example: Try yourself 5.1.2

THE NUMBER OF ELECTRONS IN A GIVEN AMOUNT OF CHARGE

The net charge on an object is $-4.8\,\mu C$ ($1\,\mu C$ = 1 microcoulomb = 10^{-6} C).
Calculate the number of extra electrons on the object.

ELECTRICAL CONDUCTORS AND INSULATORS

Electrons are much easier to move than protons. They also move more freely in some materials than in others.

In **metals**, the outermost electrons are only very slightly attracted to their respective nuclei. As a consequence, metals are good **conductors** of electricity. In conductors, loosely held electrons can effectively 'jump' from one atom to another and move freely throughout the material. This can be seen in Figure 5.1.3.

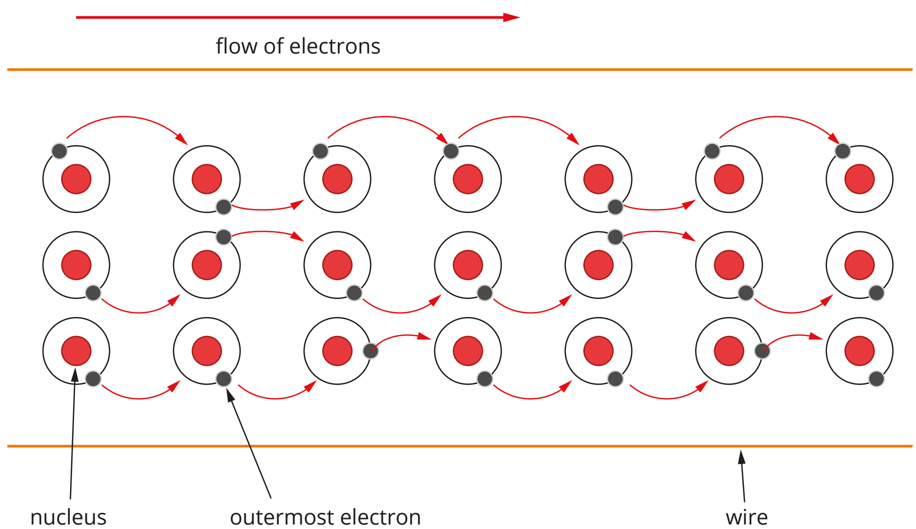


FIGURE 5.1.3 Electrons moving through a conductor. The outermost electrons are free to move throughout the lattice of positive ions.

FIGURE 5.1.4 These copper wires conduct electricity by allowing the movement of charged particles.

Copper is an example of a very good conductor. For this reason, it is used in telecommunications and electrical and electronic products (see Figure 5.1.4).

In comparison, the outer electrons in **non-metals** are very tightly bound to their respective nuclei and cannot readily move from one atom to another. Non-metals do not conduct electricity very well and are known as **insulators**. A list of common conductors and insulators is provided in Table 5.1.1.

TABLE 5.1.1 Some common conductors and insulators.

Good conductors	Good insulators
all metals, especially silver, gold, copper and aluminium and any ionic solution	plastics polystyrene dry air glass porcelain cloth (dry)
Moderate conductors	**Moderate insulators**
water earth semiconductors, e.g. silicon, germanium skin	wood paper damp air ice, snow

PHYSICS IN ACTION

Lightning

Lightning is one of nature's greatest spectacles (see Figure 5.1.5). No wonder it was so long thought of as the voice of the gods. In the mid-eighteenth century, Benjamin Franklin showed that lightning is basically the same sort of electrical phenomenon as can be achieved by rubbing a glass rod with wool, or rubbing a balloon on your hair.

FIGURE 5.1.5 Lightning bolts over a city skyline.

A typical lightning bolt transfers 10 or more coulombs of negative charge (over 60 billion billion electrons) in approximately one thousandth of a second. A moderate thundercloud with a few flashes per minute generates several hundred megawatts of electrical power, the equivalent of a small power station.

It is thought that, during a thunderstorm, charge is transferred in collisions between the tiny ice crystals that form as a result of the cooling of upwards-flowing moist air and the larger, falling hailstones. As a result of small temperature differences between the crystals and hailstones, the crystals become positively charged and the hailstones negatively charged. The crystals carry their positive charge to the top of the cloud while the negative charge accumulates in the lower region.

There is normally also a second smaller positively charged region at the bottom owing to positive charges attracted up from the ground towards the negative region, as seen in Figure 5.1.6.

There will be strong electric fields between these regions of opposite charge. If they become sufficiently strong, electrons can be stripped from the air molecules (they become ionised). Because of the electric field, the free electrons and ions will gain kinetic energy and collide with more molecules, starting an 'avalanche of charges'. This is the lightning flash seen either within the cloud or between the Earth and the cloud. Most flashes are within the cloud; only a relatively small number actually strike the ground.

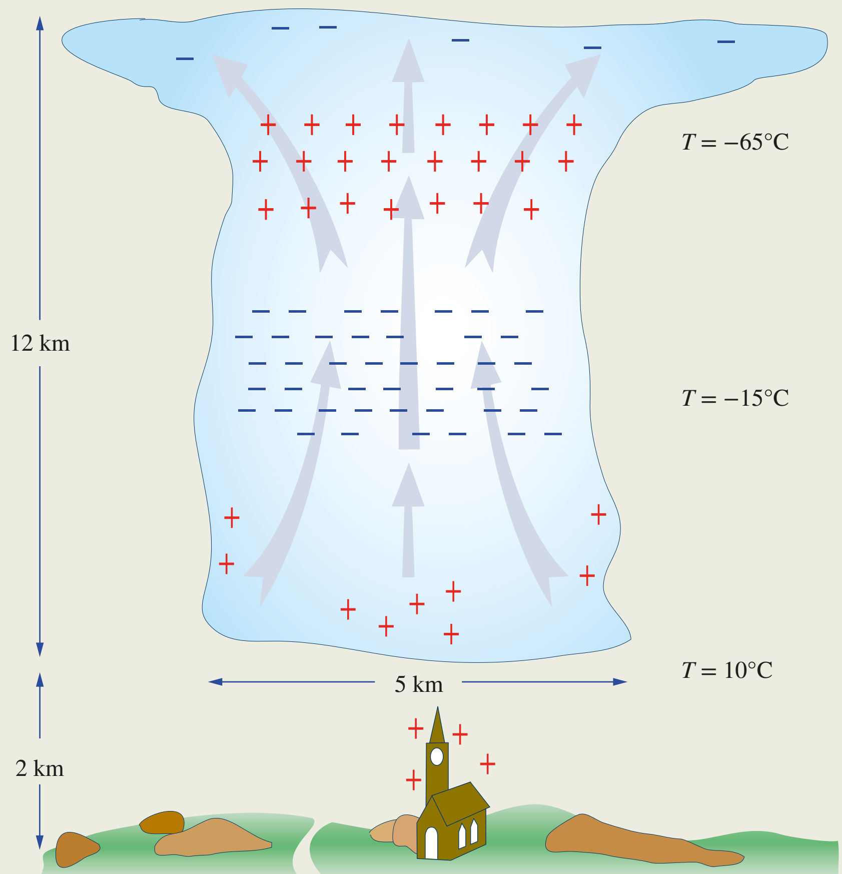


FIGURE 5.1.6 A thundercloud can be several kilometres wide and well over 10 km high. Strong updrafts drive the electrical processes that lead to the separation of charge. The strong negative charge of the lower region of the cloud will induce positive charges on tall objects on the ground. This may lead to a discharge, which can form a conductive path for lightning.

EXTENSION

Semiconductors

Some materials, such as silicon, are known as semimetals or metalloids. Their properties are somewhere between those of metals and non-metals. For example, the electrons in a silicon atom are not as tightly bound to the nucleus as those of a non-metal. However, they are not as easy to remove as the electrons in a metal. Hence, silicon and elements like it are known as semiconductors.

Silicon's ability to conduct electricity can be adjusted by adding small amounts of other elements such as boron, phosphorus, gallium, or arsenic in a process known as doping. Adding another substance contributes free electrons, which can greatly increase the conductivity of silicon within electronic devices. This makes silicon very useful in the construction of computer chips like the one shown in Figure 5.1.7.

Much of the convenience of our modern lifestyle is based on the unique conductive properties of silicon and it has many electronic applications, some of which are shown in Figure 5.1.8.

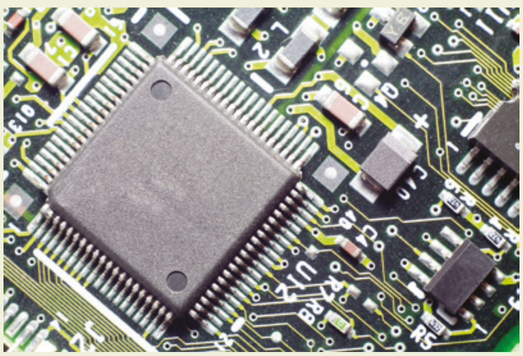


FIGURE 5.1.7 Silicon has been used in the construction of computer chips since the 1950s.

FIGURE 5.1.8 All of these electronic devices use silicon in their construction.

5.1 Review

SUMMARY

- Like charges repel; unlike charges attract.
- When an object loses electrons, it develops a positive net charge; when it gains electrons, it develops a negative net charge.
- The letter q is used to represent the quantity of charge. The SI unit of charge is the coulomb (C).
- The elementary charge (e), the charge on a proton, is equal to $+1.6 \times 10^{-19}$ C. The elementary charge, $-e$, of an electron is -1.6×10^{-19} C.
- Electrons move easily through conductors, but not through insulators. This is because the outermost electrons in materials that are good conductors are weakly attracted to the nucleus and are free to move, whereas electrons in insulators are more strongly held to the nucleus of the atom, or are involved in bonding atoms together.

KEY QUESTIONS

1. Plastic strip A, when rubbed, is found to attract plastic strip B. Strip C is found to repel strip B. What will happen when strip A and strip C are brought close together?
2. Calculate how many electrons make up a charge of -5.0 C.
3. Calculate the charge, in coulombs, of 4.2×10^{19} protons.
4. Explain why electric circuits often consist of wires that are made from copper and are coated in protective plastic.

5.2 Energy in electric circuits

In this section, the concept of **electrical potential energy** will be explored, as well as the force that causes electrons to move in a circuit.

Electrons won't move around a circuit unless they are forced to. This force is provided by the electric field created between the positive and negative terminals of a source like a battery (Figure 5.2.1). Inside every battery, a chemical reaction is taking place that moves electrons from one place and deposits them in another place. The chemical reactions convert chemical energy into potential energy as electrons are pushed from the positive terminal onto the negative terminal. This results in a separation of charge which causes an energetic electric field in the space surrounding the terminals. The potential energy stored in this electric field is known as electrical energy. When a circuit connects two ends of the battery, the electric field travels through the wires at the speed of light, carrying the potential energy with it. It is the electric field that provides the force on the electrons within the wires that causes them to move as an electric current. The electrical potential energy stored in the field is transferred by the electrons in energy transformation devices like motors or light globes and is converted into other forms of energy such as kinetic energy, heat or light.

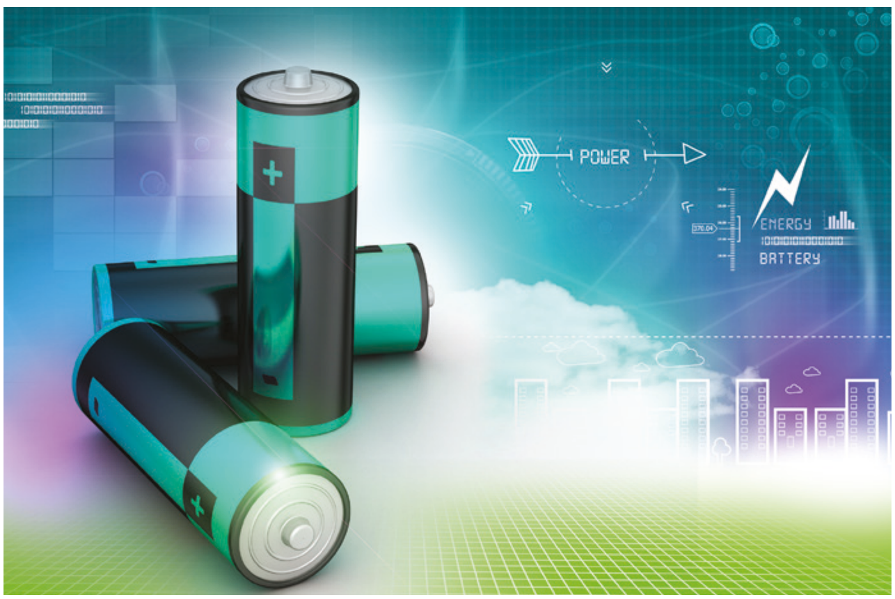


FIGURE 5.2.1 Chemical energy is stored and converted to electrical potential energy by batteries.

ENERGY IN CIRCUITS

Chemical energy is stored inside a battery as certain combinations of metals and compounds that will react in a reduction-oxidation (redox) reaction. In this type of reaction, the chemical energy is **transformed** into electrical potential energy. This potential energy is stored in the electric field that exists as a result of the separation and build-up of charge on the two terminals of the battery. This can be visualised as the coiling of invisible springs that exist between like charges that have been forced too close together. One terminal (the negative terminal) has a concentration of negative charges; the other terminal (the positive terminal) has a concentration of positive charges. When the battery is isolated (i.e. not connected in a circuit), the build-up of charge will oppose the redox reaction and so the reaction stops and the unused battery can maintain its reactants for years. Once the battery is connected to a device by wires, the redox reactions will continue for some time and will maintain the difference in charge between the two terminals.

The difference in electrical potential energy between the two terminals of a battery can be quantified as a difference in the electrical potential energy per unit charge. This is commonly called **potential difference** (ΔV) and is measured in **volts** (V).

PHYSICSFILE

Cells and batteries

A single cell generates electricity by converting chemical energy to electrical potential energy. A commercial battery can be a single cell or a multiple of cells connected together. Often a series of cells are packaged in a way that makes it look like a single device, but inside is a battery of cells connected together (see Figure 5.2.2). The terms 'battery' and 'cell' can be used interchangeably as the term 'battery' is frequently used in common language to describe a cell.

FIGURE 5.2.2 A mobile-phone battery. The term battery actually refers to a group of electric cells connected together.

EXTENSION

Fields

Field theory says that there are certain properties of matter that create regions around an object in which matter with the corresponding property will experience a force. These forces are usually described as 'acting at a distance'. Mass creates a gravitational field in which other masses will experience an attractive force called gravity. Similarly, the poles of a magnet will create a field that will either attract or repel the pole of another magnet.

Electric fields form around any charged object. In these fields, other objects of charge will experience attractive or repulsive forces. Electric fields can be represented by arrows that show the direction that a small positive charge would move if it were placed in the field. The direction of the field would be away from a positively charged plate towards a negatively charged plate. An electron placed in an electric field would experience a force that is away from the negative plate and towards the positive plate, which is in the opposite direction of the field direction. This is shown in Figure 5.2.3.

The shape of the field around some charged objects is shown in Figure 5.2.4. The concept of fields will be developed further in Unit 3 Year 12.

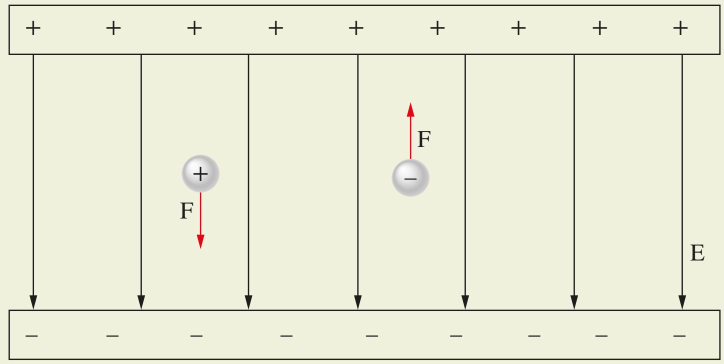


FIGURE 5.2.3 An electric field is created between two charged plates. The direction of the field is shown by the arrows labelled *E*, and the force on a positive particle and a negative particle is represented by the force vectors labelled *F*.

Electric and magnetic fields are actually unified into one type of field called an electromagnetic field. This implies that there is a connection between magnetism and electricity. One feature of electromagnetic fields is that they spread at the speed of light.

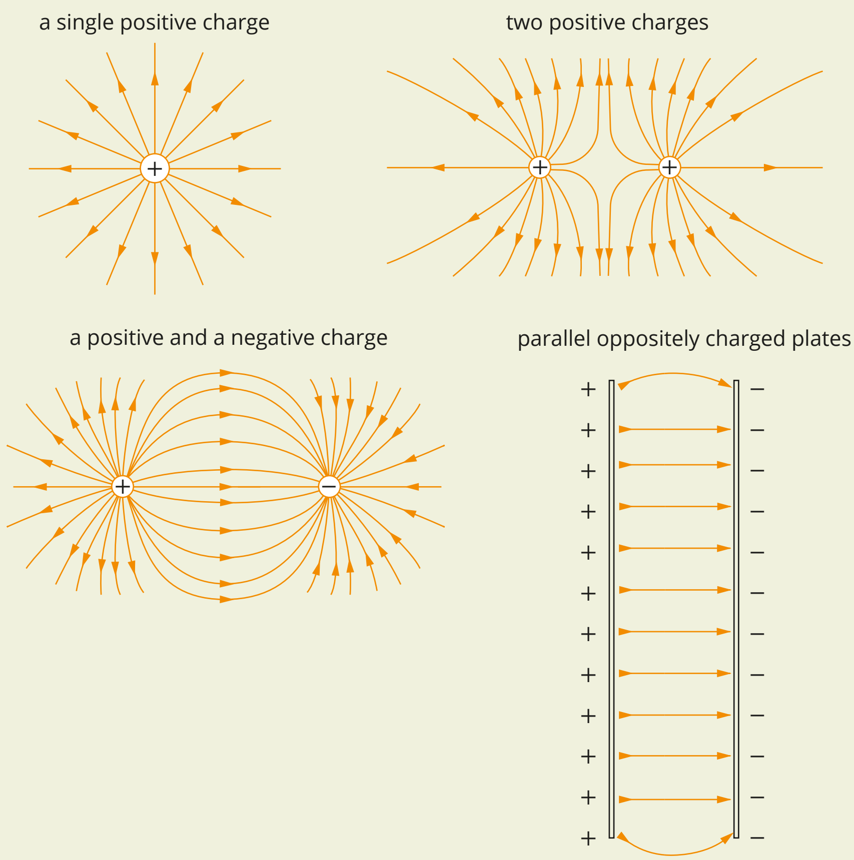


FIGURE 5.2.4 Electric fields can be represented by lines showing the direction of the field. The closeness of the lines indicate the strength of the field.

It is this potential difference at the terminals of the battery that provides the energy to a circuit. The energy is stored in the electric field and then transferred via electrons to different components in the circuit. At each component the energy is transformed into a different type of energy. For example, the energy could be transformed into heat and light if the component is a light bulb. If the component is a fan, the energy is transformed into kinetic energy (motion) and some heat and sound.

Energy transfers and transformations in a torch

A torch is a simple example of how energy is transformed and transferred within a circuit. In the torch shown in Figure 5.2.5, chemical energy in the battery is transformed into electrical potential energy in the electric field. There are two batteries connected in series so a bigger potential difference is available. Energy can be transferred via the electrons to the light bulb once the end terminals of the batteries have been connected to the torch's circuit: that is, when the torch is switched on. Once the connection is made in the switch, the electric field spreads through the wire at the speed of light. The electric field applies a force on every charged object within the wire—protons in the nucleus and electrons around the atom. The positively charged nuclei are held in place by the lattice structure of metals, while the inner electrons are held onto strongly by their attraction to the nucleus, and so they will not move. The delocalised outer electrons, however, are only loosely held and so they will move in the opposite direction to the electric field. As they move, they interact with the matter in the device and transfer the energy from the field into other forms. In the case of a torch they transform electrical potential energy in the electric field into the kinetic energy of the atoms within the filament wire in the bulb, which is then emitted as heat and light.

The energy changes can be summarised as:

$$\text{chemical energy} \xrightarrow{\text{transformed}} \text{electrical potential energy}$$

$$\text{electrical potential energy} \xrightarrow{\text{transformed}} \text{kinetic energy (filament atoms)}$$

$$\text{kinetic energy (filament atoms)} \xrightarrow{\text{transformed}} \text{heat energy + light}$$

Eventually, when most of the chemicals within the battery have reacted, the battery is no longer able to provide enough electrical potential energy to power the torch. This is because the chemical reaction has slowed and electrons are not being driven to the negative terminal in sufficient numbers. The torch stops working and the batteries are said to have gone flat.

Quantifying potential difference

As with other forms of energy, it is useful to be able to quantify the amount of potential difference in a given situation. Potential difference is formally defined as the amount of electrical potential energy given to each coulomb of charge. As an equation, it is:

$$\Delta V = \frac{E}{q}$$

where ΔV is potential difference (V)

E is electric potential energy (J)

q is charge (C).

Since energy is measured in joules and charge in coulombs, the potential difference is measured in joules per coulomb (J C^{-1}). This quantity has been assigned a unit, the volt (V) to honour Alessandro Volta, who invented the first battery. A potential difference of $1\,\text{J C}^{-1}$ is equal to 1 V. When a battery is labelled 9 V, this means that the battery provides 9 joules of energy to each coulomb of charge.

PHYSICSFILE

Volts and potential difference

Somewhat confusingly, the formula and data sheet uses the symbol 'V' for both the quantity potential difference and its unit of measurement, the volt. In this text, for the quantity of potential difference, we will use the symbol ΔV (Delta V). For the quantity of electrical potential, the symbol V is used. The unit, volts, has the symbol V, which is not in italics; the context usually makes it clear which meaning is intended. Potential difference is sometimes referred to as 'voltage'; however, in this text the more correct term will be used.

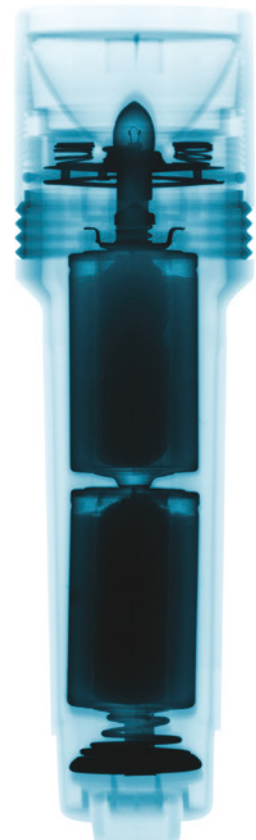


FIGURE 5.2.5 An X-ray image of the internal structure of a torch. The bulb and two batteries are clearly visible.

PHYSICSFILE

Birds on a wire

Birds can sit on power lines and not get electrocuted even though the wires are not insulated.

For a current to flow through a bird on a wire, there would have to be a potential difference between its two feet. Since the bird has both feet touching the same wire, which might be at a very high potential (voltage), there is no potential *difference* between the bird's feet. If the bird could stand on the wire and touch any other object such as the ground or another wire, then it would get a big electric shock.

This is because there would be a potential difference between the wire and the other object causing a current to flow.

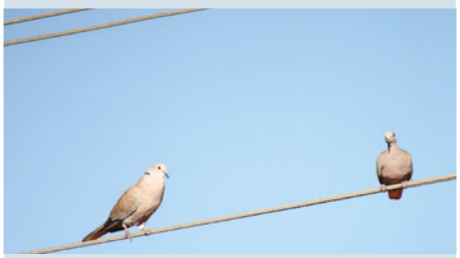


FIGURE 5.2.6 There is no potential difference between each bird's feet.

Worked example 5.2.1

DEFINITION OF POTENTIAL DIFFERENCE

Calculate the amount of electrical potential energy carried by 5 C of charge at a potential difference of 10 V.	
Thinking	**Working**
Recall the definition of potential difference.	$\Delta V = \frac{E}{q}$
Rearrange this to make energy the subject.	$E = \Delta Vq$
Substitute in the appropriate values and solve.	$E = 10 \times 5$ $= 50\,\text{J}$

Worked example: Try yourself 5.2.1

DEFINITION OF POTENTIAL DIFFERENCE

A car battery can provide 3600 C of charge at 12 V. How much electrical potential energy is stored in the battery?

Measuring potential difference: the voltmeter

Potential difference is usually measured by a device called a **voltmeter**.

Voltmeters are wired into circuits to measure the change in potential as current passes through a particular component. This means that one wire of the voltmeter is connected to the circuit before the component and the other wire is connected to the circuit after the component. This is called connecting the voltmeter 'in parallel'.

In Figure 5.2.7, the voltmeter is connected to the circuit either side of the light globe. This is so that it can measure the potential difference (voltage drop) across the light globe. It is important to connect the voltmeter with the positive terminal closest to the positive terminal of the power supply. The voltmeter's negative terminal is connected closest to the negative terminal of the power supply.

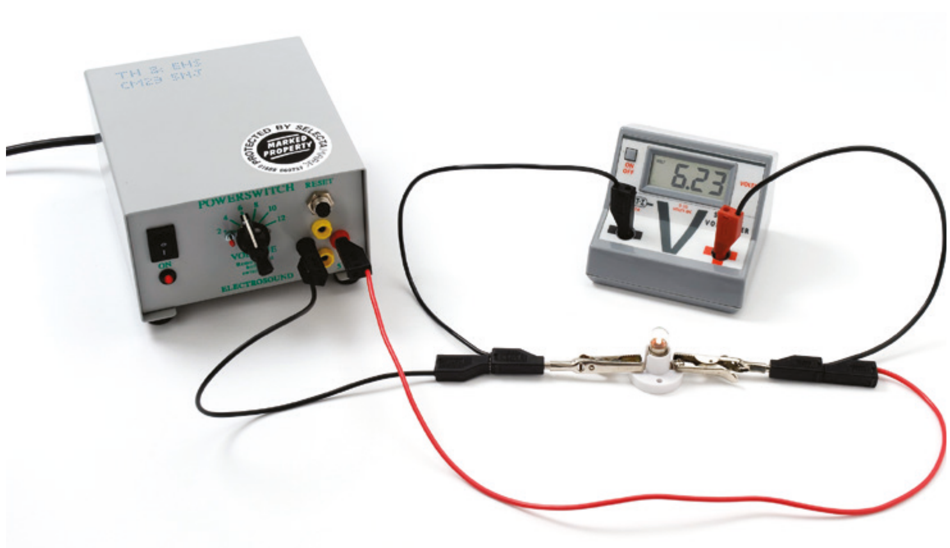


FIGURE 5.2.7 A voltmeter measures the potential difference (in this case, 6.23 V) across a light globe.

ANALOGIES FOR POTENTIAL DIFFERENCE

Analogies help us to understand concepts that cannot be seen directly. If you can analyse the way in which the analogy functions then you can make comparisons to the concept you are studying. There are a number of popular analogies for electrical potential difference, including pumps that cause water pressure in a pipe and pedals pushing a bike chain around the cogs. Perhaps the best analogy compares a source of electrical potential energy to an escalator at a shopping mall. Just as a dry cell uses energy to push electrons to a higher energy state, so an escalator uses energy to lift people from the ground level to a higher level in a shopping mall, where they have greater gravitational potential energy. This analogy neatly ties the electrical potential energy gain in a dry cell to the gravitational potential energy gain in the mall, as in both cases the energy is stored in a field ready to do work via the electrons or the people.

5.2 Review

SUMMARY

- Electric potential difference measures the difference in electric potential energy available per unit charge
- Potential difference can be defined as the work done to move a charge against an electric field between two points, using the equation:

  $E = \Delta V q$ or $\Delta V = \dfrac{E}{q}$
- In a circuit, the energy required for charge separation is provided by a cell or battery. The chemical energy within the cell is transformed into electric potential energy stored in an electric field.

KEY QUESTIONS

1 Under what conditions will charge flow between two bodies linked with a rod? Choose the correct response from the following options.
   A The potential difference between the bodies is not zero and the rod is made of a conducting material.
   B The potential difference between the bodies is not zero and the rod is made of an insulating material.
   C The potential difference between the bodies is equal to zero and the rod is made of a conducting material.
   D The potential difference between the bodies is equal to zero and the rod is made of an insulating material.

2 Calculate the potential of a battery that gives a charge of 10 C:
   a 40 J of energy
   b 15 J of energy
   c 20 J of energy.

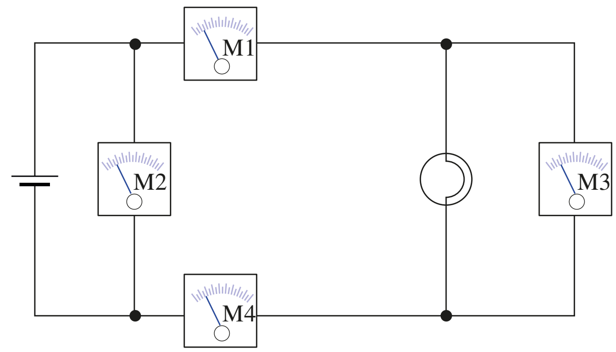
3 A charge of 5 C flows from a battery through an electric water heater and delivers 100 J of heat to the water. What was the potential difference of the battery?

4 How much charge must have flowed through a 12 V car battery if 2 kJ of energy was delivered to the starter motor?

5 A light globe that is connected to 240 V uses 3.6 kJ of electric potential energy in one minute.
   a Into what type(s) of energy has the electrical energy been transformed?
   b Calculate the charge that flows through the light globe.

6 The electrical energy obtained from a battery can be compared to the energy of water stored in a hydroelectric dam in the mountains. In this model, to what could the potential difference of the battery be likened?

7 Andrea wishes to measure the potential difference for a light globe. She has set up a circuit as shown in the figure below.



In which positions (M1, M2, M3 or M4) can she place a voltmeter?

5.3 Electric current and circuits

A flow of electric charge is called electric current. Current can be carried by moving electrons in a wire or by ions in solution. This section explores current as it flows through wires in electric circuits.

Electric circuits are involved in much of the technology used every day and are responsible for many familiar sights (Figure 5.3.1). To construct electric circuits, you must know about the components of a circuit and be able to read circuit diagrams.

FIGURE 5.3.1 Electric circuits are responsible for lighting up whole cities.

ELECTRIC CIRCUITS

An **electric circuit** is a path made of conductive material, through which charges can flow in a closed loop. This flow of charges is called **electric current**. The most common conductors used in circuits are metals, such as copper wire. The charges that flow around the circuit within the wire are negatively charged delocalised electrons. The movement of electrons in the wire is called **electron flow**.

Electric fields created by the separation of charge in a cell can travel through air, but air is a very poor conductor of electricity and so no electrons will leave the negative surface and travel to the positive surface if the potential difference is small and the air gap is large. If, however, a substance like copper is placed in between the two charged ends of a cell, the electric field will concentrate and travel through the metal in preference to spreading through the air. The electric field spreads through the circuit at the speed of light and causes delocalised electrons to move in the opposite direction to the electric field. It is interesting to note that while the electric field propagates through the circuit at the speed of light, the free electrons drift through the circuit at the speed of snails, that is millimetres per minute.

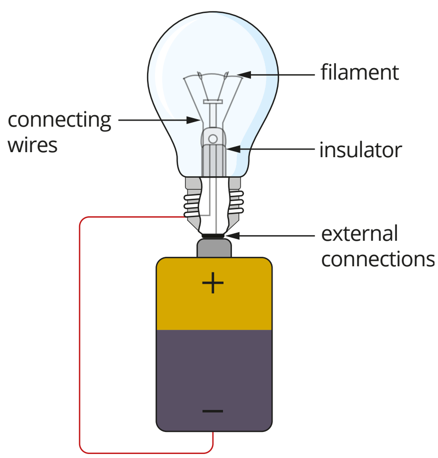


FIGURE 5.3.2 When there is a complete conduction path from the positive terminal of a battery to the negative terminal, a current flows.

A simple example of an electric circuit is shown in Figure 5.3.2. The light bulb is in contact with the positive terminal of the battery; a copper wire joins the negative terminal of the battery to one end of the filament in the light bulb. This arrangement forms a closed loop that allows electrons within the circuit to flow away from the negative terminal towards the positive terminal of the battery. The battery is a source of energy and provides the electric field that spreads through the circuit. The light bulb converts this energy into heat and light energy when the circuit is connected.

If a switch is added to the circuit in Figure 5.3.2, the light bulb can be turned off and on. When the switch is closed, the electric field spreads through the wires and the circuit is complete. The current flows in a loop, in response to the electric field, along a path made by the conductors.

When the switch is open, there is a break in the circuit, which means that there is an air gap between the terminals of the switch. Although the field is established through the air gap, no current can flow through the gap if the potential difference is low enough and the separation is large enough, therefore current can no longer flow. This is what happens when you turn off the switch for a lamp or TV. A circuit where the conducting path is broken is often called an open circuit.

One common misconception about current is that charges are used up or lost when a current flows around a circuit. However, the charge-carrying electrons are conserved at all points in a circuit.

> Current will flow in a circuit only when the circuit forms a continuous (closed) loop from one terminal of a power supply to the other terminal.

REPRESENTING ELECTRIC CIRCUITS

Common symbols for electronic components

A number of different components can be added to a circuit. It is not necessary to be able to draw detailed pictures of these components; simple symbols are much clearer. The common symbols used to represent the electrical components in electric circuits are shown in Figure 5.3.3.

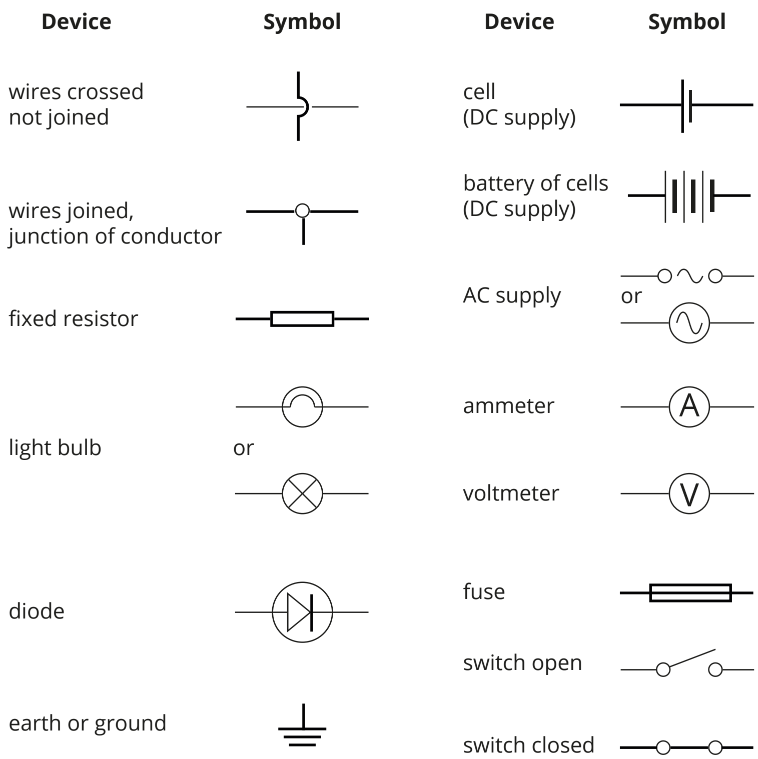


FIGURE 5.3.3 Some commonly used electrical devices and their symbols.

Circuit diagrams

When building anything, it is important that the builder has a clear set of instructions from the designer. This is as much the case for electric circuits as it is for a tall building or a motor vehicle.

Circuit diagrams are used to clearly show how the components of an electric circuit are connected. They simplify the physical layout of the circuit into a diagram that is recognisable by anyone who knows how to interpret it. You can use the list of common symbols for electrical components (Figure 5.3.3) to interpret any circuit diagrams.

The circuit diagram in Figure 5.3.4b shows how the components of the torch shown in Figure 5.3.4a are connected in a circuit. The circuit can be traced by following the straight lines representing the connecting wires.

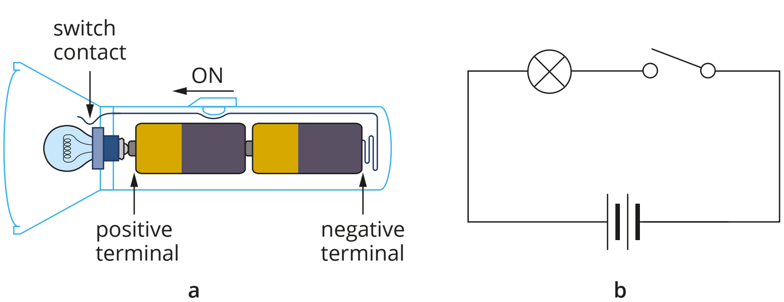


FIGURE 5.3.4 (a) A battery and light bulb connected by conductors in a torch constitute an electric circuit. (b) The torch's circuit can be represented by a simple circuit diagram.

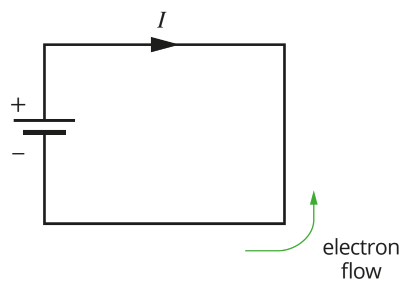


FIGURE 5.3.5 Conventional current (I) and electron flow are in opposite directions. The long terminal of the battery is positive.

Conventional current vs electron flow

When electric currents were first studied, it was incorrectly thought the charges that flowed in circuits were positive. Based on this, scientists traditionally talked about electric current as if current flowed from the positive terminal of the battery to the negative terminal. This convention is still used today, even though it is now known that it is actually the negative charges (electrons) that flow around a circuit.

On a circuit diagram current is indicated by a small arrow and the symbol I. This is called **conventional current** or just **current**. The direction of conventional current is opposite to the direction of electron flow (Figure 5.3.5).

Conventional current (or current), I, flows from the positive terminal of a power supply to the negative terminal.

Electron flow (or electron current) refers to the flow of electrons from the negative terminal to the positive terminal of a power supply.

The amount of charge is the same in both cases for the same circuit.

QUANTIFYING CURRENT

In common electrical circuits, a current consists of electrons flowing within a copper wire (Figure 5.3.6). This current, I, can be defined as the amount of charge that passes through a point in the conductor per second.

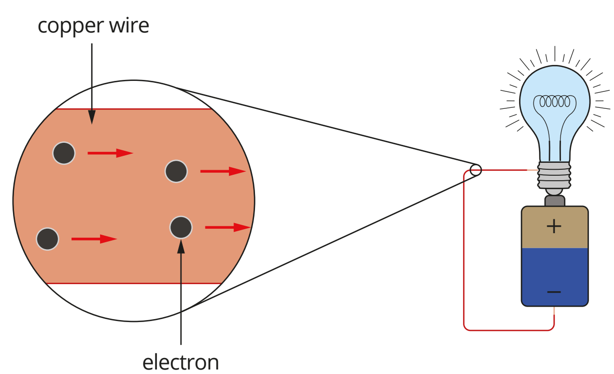


FIGURE 5.3.6 The number of electrons that pass through a point per second gives a measure of the current. Because the electrons do not leave the wire, current is conserved in all parts of the circuit.

An equation to express this is:

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

where I is the current in amperes

q is the amount of charge in coulombs

t is the number of seconds that have passed.

Current is measured in amperes, or amps (A). One ampere is equivalent to one coulomb per second ($C\,s^{-1}$).

Current is a flow of charge, but the charge is carried by electrons. The charge that flows is equal to the number of electrons (n_e) that flow through a particular point in the circuit multiplied by the charge on one electron ($q_e = -1.6 \times 10^{-19}$ C). When this total charge is divided by the time that has elapsed in seconds (t) this gives us the current in amperes (A). This makes the equation:

$$I = \frac{q}{t} = \frac{n_e q_e}{t}$$

A typical current in a circuit powering a small DC motor would be about 50 mA. Even with this seemingly small current, approximately 3×10^{17} electrons flow past any point on the wire each second.

Worked example 5.3.1

USING $I = \frac{q}{t}$

Calculate the number of electrons that flow past a particular point each second in a circuit that carries a current of 0.5 A.	
Thinking	**Working**
Rearrange the equation $I = \frac{q}{t}$ to make q the subject.	$I = \frac{q}{t}$ so $q = I \times t$
Calculate the amount of charge that flows past the point in question by substituting the values given.	$q = 0.5 \times 1$ $= 0.5\,\text{C}$
Find the number of electrons by dividing the charge by the charge on an electron (1.6×10^{-19} C).	$n_e = \frac{q}{q_e}$ $= \frac{0.5}{1.6 \times 10^{-19}}$ $= 3.12 \times 10^{18}$ electrons

Worked example: Try yourself 5.3.1

USING $I = \frac{q}{t}$

Calculate the number of electrons that flow past a particular point each second in a circuit that carries a current of 0.75 A.

MEASURING CURRENT: THE AMMETER

Current is commonly measured by a device called an **ammeter** (see Figure 5.3.7).

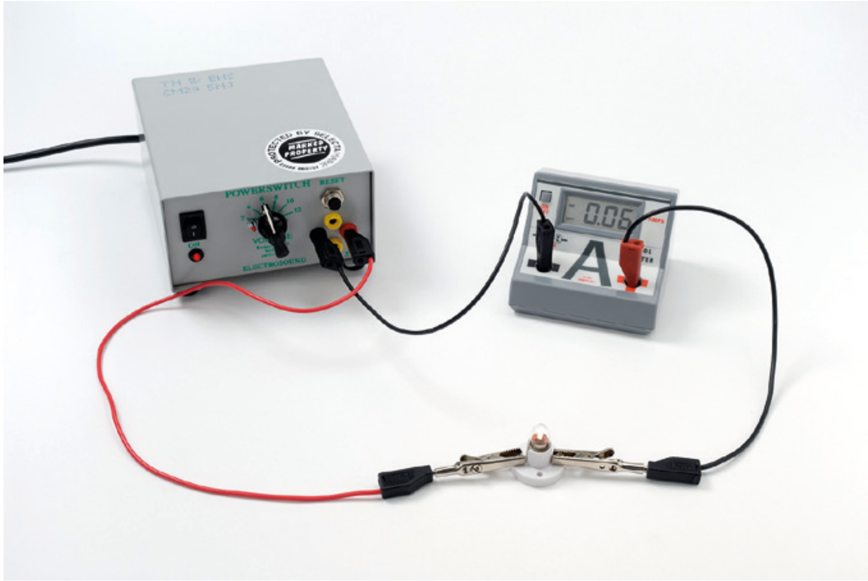


FIGURE 5.3.7 A digital ammeter (labelled with an A) measures current in a circuit.

Figure 5.3.7 shows the ammeter connected along the same path taken by the current flowing through the light bulb. This is referred to as connecting the ammeter 'in series'. Series circuits are covered in more detail later in this chapter. The positive terminal of the ammeter is connected so that it is closest to the positive terminal of the power supply. The negative terminal of the ammeter is closest to the negative terminal of the power supply.

PHYSICSFILE

Multimeters

The internal circuitry of a voltmeter is significantly different to an ammeter. Electricians and scientists who work with electrical circuits often find it inconvenient to keep collections of voltmeters and ammeters so they have found a way to bundle the circuitry for both meters up into a single device known as a *multimeter*. A multimeter is shown in Figure 5.3.8.

This is much more convenient because the multimeter can be quickly switched from being an ammeter to being a voltmeter as needed. However, when a multimeter is switched from one mode to another, it is important to make a corresponding change to the way it is connected to the circuit being measured. An ammeter is connected in series and a voltmeter in parallel. In fact, if a multimeter is working as an ammeter and it is connected in parallel like a voltmeter, it may draw so much current that its internal circuitry will be burnt out and the multimeter will be destroyed.

FIGURE 5.3.8 A digital multimeter can be used as either an ammeter or a voltmeter.

Measuring the current is possible because charge is conserved at all points in a circuit. This means that the current that flows into a light bulb is the same as the current that flows out of the light bulb. An ammeter can therefore be connected before or after the bulb in series to measure the current. Table 5.3.1 lists some typical values for electric current in common situations.

TABLE 5.3.1 Typical values for electric current.

Situation	Current
lightning	10 000 A
starter motor in car	200 A
fan heater	10 A
toaster	3 A
light bulb	400 mA
pocket calculator	5 mA
nerve fibres in body	1 μA

WORK DONE BY A CURRENT IN A CIRCUIT

In electrical circuits, electrical potential energy is converted into other forms of energy. When energy is changed from one form to another, work is done. (Work is covered in more detail in Chapter 9.) The amount of energy provided by a circuit can be calculated using the definitions for potential difference $\Delta V = \frac{E}{q}$ and current $I = \frac{q}{t}$.

Rearranging the definition of potential difference gives:

$E = \Delta Vq$

Using the definition of current:

$q = It$

Therefore:

$E = \Delta VIt$

where E is the energy provided by the current, which is the same as the work done (J)

ΔV is the potential difference (V)

I is the current (A)

t is the time (s).

This gives us a practical way to calculate the energy used in a circuit from measurements we can make.

Worked example 5.3.2

USING $E = \Delta VIt$

A potential difference of 12 V is used to generate a current of 750 mA to heat water for 5 minutes. Calculate the energy transferred to the water in that time.	
Thinking	**Working**
Convert the quantities to SI units.	$\frac{750\text{ mA}}{1000} = 0.750\text{ A}$ $5\text{ min} \times 60\text{ s} = 300\text{ s}$
Substitute values into the equation and calculate the amount of energy in joules.	$E = \Delta VIt$ $= 12 \times 0.750 \times 300$ $= 2700\text{ J}$

Worked example: Try yourself 5.3.2

USING $E = \Delta VIt$

A potential difference of 12 V is used to generate a current of 1750 mA to heat water for 7.5 minutes. Calculate the energy transferred to the water in that time.

Rate of doing work: power

If you wanted to buy a new kettle, you might wonder how you could determine how quickly different kettles boil water.

Printed on all appliances is a rating for the power of that device. **Power** is a measure of how fast energy is converted by the appliance. In other words, power is the rate at which energy (E) is transformed by the components within the device. This can also be described as the rate at which work (W_d) is done. As an equation:

$$P = \frac{\text{work done}}{\text{time}} = \frac{W_d}{t} = \frac{E_{\text{transformed}}}{t} = \frac{\Delta VIt}{t} = \Delta VI$$

where P is the power in joules per second (J s^{-1}). One joule per second is 1 watt (W).

The more powerful an appliance is, the faster it can do a given amount of work. In other words, an appliance that draws more power can do the same amount of work in a shorter amount of time. If you want something done quickly, then you need an appliance that has a higher power rating.

Rearranging the expression above enables us to calculate the energy transformations in a circuit by measuring potential difference and current across circuit components and the time for which the current flows.

Given $P = \frac{W_d}{t} = \Delta VI$

Then work done, $W_d = \Delta VIt$.

The power dissipated by those components can be calculated in watts (W), and the work done is given in joules (J).

Worked example 5.3.3

USING $P = \Delta VI$

An appliance running on 230 V draws a current of 4 A. Calculate the power used by this appliance.	
Thinking	**Working**
Identify the relationship needed to solve the problem.	$P = \Delta VI$
Identify the required values from the question, substitute and calculate.	$P = 230 \times 4$ $= 920$ W

Worked example: Try yourself 5.3.3

USING $P = \Delta VI$

An appliance running on 120 V draws a current of 6 A. Calculate the power used by this appliance.

ANALOGIES FOR ELECTRIC CURRENT

Since we cannot see the movement of electrons in a wire, it is sometimes helpful to use analogies or 'models' to visualise or explain the way an electric current behaves. It is important to remember that no analogy is perfect, so there may be situations where the electric current does not act as you would expect from the analogy.

Water model

A very common model is to think of electric charges as water being pumped around a pipe system that is already full of water, as shown in Figure 5.3.9. The electric field created by the battery pushes electrons through the wires just like the pressure created by a pump pushes water through the pipes. Since water cannot disappear, the same amount of water flows in every part of a pipe, just as the electric current is the same in every part of a wire. Light bulbs in an electric circuit are like turbines: whereas the turbine converts the energy of the flowing water into the kinetic energy of the turbine, a light bulb converts electrical energy into heat and light. The water that has flowed through the turbine flows back to the pump to be forced back into the pipes, just as the electrons eventually flow back to the battery to be pushed towards the negative terminal and into the circuit.

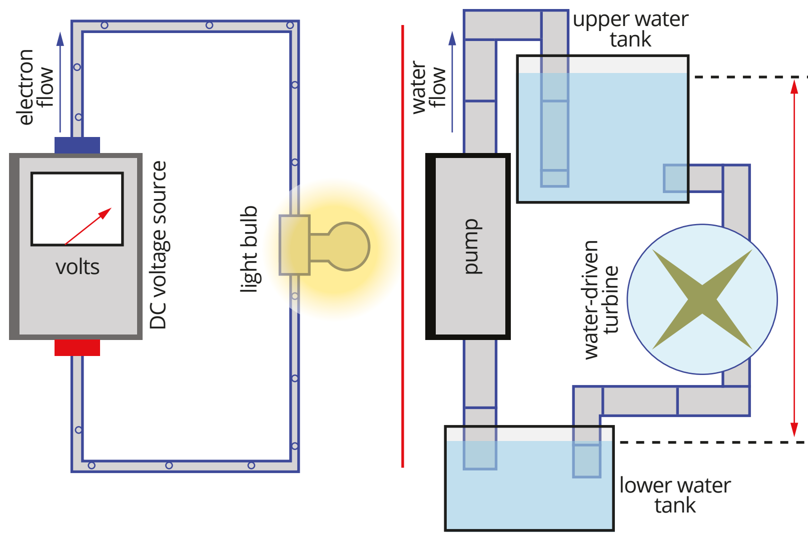


FIGURE 5.3.9 An electric current can be compared to water flowing through a pipe system.

This model explains the energy within a circuit quite well.

- The power supply transfers energy to the electric field by forcing the electrons towards the negative end of the battery.
- The electrical energy in the electric field is transferred via the electrons in the components in the circuit and is converted into other forms.

5.3 Review

SUMMARY

- Current will flow in a circuit only when the circuit forms a continuous (closed) loop from one terminal of a power supply to the other terminal.
- When an electric current flows, free electrons all around the circuit move towards the positive terminal, at the same time due to the electric field spreading at the speed of light. This is called electron flow.
- Conventional current in a circuit flows from the positive terminal to the negative terminal.
- Current, I, is defined as the amount of charge, q, that passes through a point in a conductor per second. It has the unit amperes or amps (A), which are equivalent to coulombs per second. The equation for this is:

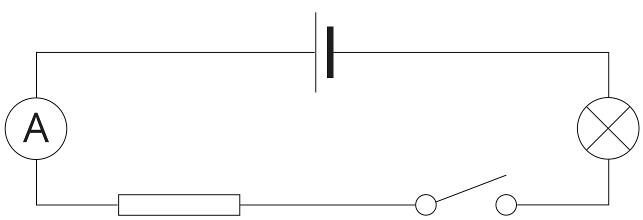
  $I = \frac{q}{t} = \frac{n_e q_e}{t}$
- Current is measured with an ammeter connected along the same path as the current flowing (in series) within the circuit.
- Power is the rate at which energy is transformed in a circuit component. It is defined and quantified by the relationships:

  $P = \frac{q}{t} = \Delta VI$

KEY QUESTIONS

1 What are the requirements for current to flow in a circuit?

2 List the electronic devices shown in the circuit diagram below.



3 Why do scientists refer to conventional current as flowing from positive to negative?

A Protons flow from the positive terminal of a battery to the negative terminal.

B Electrons flow from the positive terminal of a battery to the negative terminal.

C Originally, scientists thought charge carriers were positive.

D Charges flow in both directions in a wire; conventional current refers to just one of the flows.

4 Calculate the current flowing in a light bulb through which a charge of 30 C flows in:

a 10 seconds?

b 1 minute?

c 1 hour?

5 A car headlight may draw a current of 5 A. Calculate the charge flowing through it in:

a 1 second?

b 1 minute?

c 1 hour?

6 Using the values given in Table 5.3.1, calculate the amount of charge that would flow through a:

a pocket calculator in 10 min

b car starter motor in 5 s

c light bulb in 1 h.

7 10^{20} electrons flow past a point in 4 seconds. Calculate:

a the amount of charge, in coulombs, that moves past a point in this time

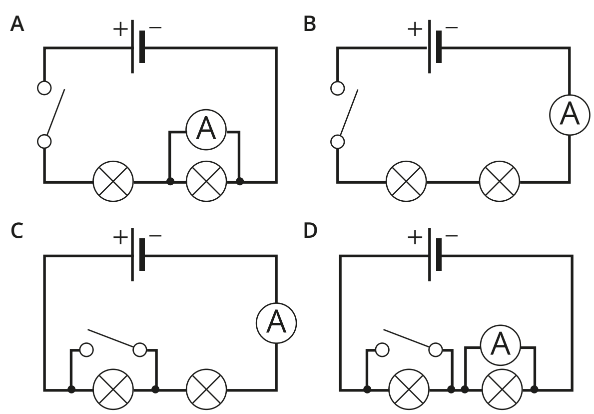
b the current, in amps.

8 3.2 C flow past a point in 10 seconds. Calculate:

a the number of electrons that move past a point in this time

b the current, in amps.

9 Which of the circuits shown in the figure below would enable you to measure the current passing through both light bulbs when the switch is closed?



10 A freezer has a power rating of 460 W and it is designed to be connected to 230 V. Calculate:

a the work performed by the freezer in 5 minutes

b the current flowing through the freezer.

5.4 Resistance

Resistance is an important concept because it links the ideas of potential difference and current. **Resistance** is a measure of how hard it is for current to flow through a particular material. As conductors allow current to pass through easily, they are said to have low resistance. Insulators have a high resistance because they 'resist' or limit the flow of charges through them.

For a particular object or material, the amount of resistance can be quantified (given a numerical value). This means that the performance of electrical circuits can be studied and predicted with a high degree of confidence.

- Resistance is a measure of how hard it is for current to flow through a particular material.
- The unit for resistance is ohms (Ω).

RESISTANCE TO THE FLOW OF CHARGE

Energy is required to create and maintain an electric current. For electrons to move from one place to another, they need to first be separated from their atoms and then given energy to move. In some materials (i.e. conductors), the amount of energy required for this is negligible (almost zero). In insulators, a much larger amount of energy is required.

Once the electrons are moving through the material, energy is also required to keep them moving at a constant speed. Consider an electron travelling through a piece of copper wire. It is common to imagine the wire as an empty pipe or hose through which electrons flow. However, a piece of copper wire is not empty—it is full of copper ions. These ions are packed tightly together in a lattice arrangement. As an electron moves through the wire, it will 'bump' into the ions. The electron needs constant 'energy boosts' to keep it moving in the right direction. This is why an electrical device will stop working as soon as the energy source (e.g. battery) is disconnected.

PHYSICSFILE

Electron movement

Even when current is not flowing, free electrons tend to move around a piece of metal due to thermal effects. The free electrons are rushing around at random with great speed. The net speed of an electron through a wire, however, is quite slow. Figure 5.4.1 compares the random motion of an electron when current is not flowing (AB) to the motion of the electron when current is flowing (AB′). The difference between the two paths is only small. However, the combined effect of countless electrons moving together in this way represents a significant net movement of charge.

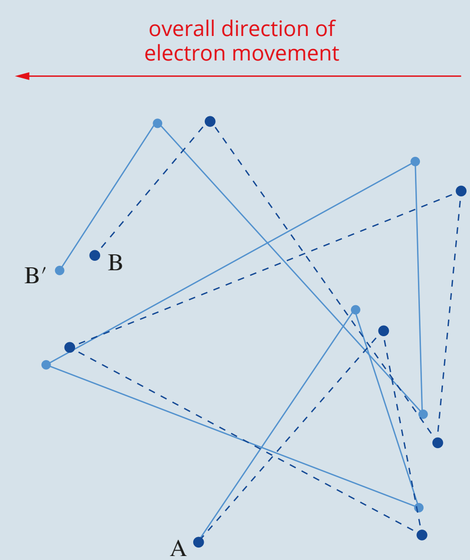


FIGURE 5.4.1 Path AB shows the random motion of an electron due to thermal effects. Path AB' shows the path of the same electron when an electric current is flowing in the direction indicated.

EXTENSION

Variables that affect resistance

Effect of cross-sectional area and length on resistance

Understanding the way electrons move through a wire can help us make some predictions about the resistance of different objects.

For example, in a longer piece of wire, the electrons bump into more ions along the way, so more energy would be needed for the electrons to travel from one end to the other. In other words, a longer piece of wire would provide greater 'resistance' to the flow of electric current.

Similarly, a thicker piece of wire allows more electrons to flow through it at the same time, much like a dual-lane highway allows faster traffic flow than a single lane. In practice, the cross-sectional area of the wire (its area when viewed end on) is important. The greater the cross-sectional area of the wire, the lower its resistance will be.

Calculating the effect of length and area on resistance

The relationship between the resistance of a conductor and its length and thickness follows a mathematical relationship. There is a direct relationship between resistance and length: doubling the length of the conductor doubles its resistance. There is an inverse relationship between resistance and the cross-sectional area of the conductor. These relationships are captured in the equation:

$$R = \frac{\rho L}{A}$$

where R is resistance (Ω), L is length (m), A is cross-sectional area (m^2) and ρ is resistivity (Ω m), a property of the material from which the conductor is made.

Temperature and resistance

Another factor that affects the resistance of a material is its temperature. The temperature of an object is a measure of the average kinetic energy of its particles. The temperature of a solid is an indication of how quickly its particles are vibrating.

Increasing the temperature of a piece of copper wire means that the copper ions will vibrate back and forth more quickly. This makes it more likely that an electron will collide with the ion as it moves past it. Therefore, increasing the temperature of the wire also increases the resistance of the wire.

Similarly, current passing through a conductor can cause it to heat up. Think again of an electron moving through a copper wire: when the electron collides with a copper ion, it loses some of its kinetic energy. However, due to this collision, the copper ion gains kinetic energy, causing it to vibrate more quickly. An increase in the kinetic energy of the copper means that its temperature has increased, so the copper wire heats up.

This is one of the reasons why personal computers contain cooling fans, as shown in Figure 5.4.2. Electrical components are packed very tightly together on the computer motherboard. Cooling the components and the conductors that connect them prevents the computer from overheating. It also reduces the resistance of the components and helps them to run more efficiently.

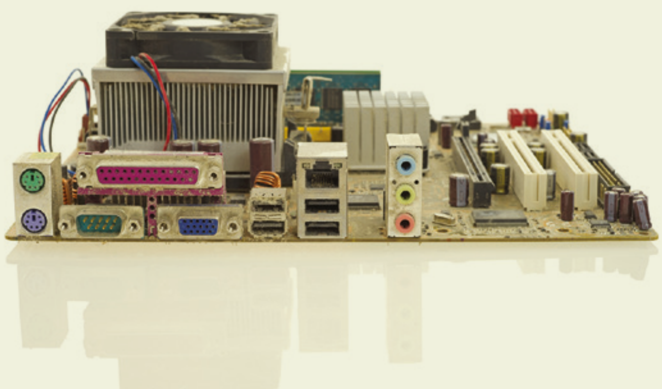


FIGURE 5.4.2 The cooling fan in this computer motherboard circulates air around the electrical components to cool them down.

Mathematically, the relationship is expressed as follows:

$$R = R_0[1 + \alpha(T - T_0)]$$

where R is the resistance of the conductor (Ω) at temperature T(K), R_0 is the resistance (Ω) at temperature T_0(K) and α is the temperature coefficient of resistance (K^{-1}) (a property of each material which describes how big an impact changing temperature has on the resistance of the material). Temperature coefficient of resistance values are shown for some common conducting materials in Table 5.4.1.

TABLE 5.4.1 Temperature coefficient of resistance values for some common conducting materials.

Substance	copper	tungsten	nickel	iron	steel
α (K^{-1})	3.9×10^{-3}	4.5×10^{-3}	6.0×10^{-3}	5.0×10^{-3}	3.0×10^{-3}

PHYSICS IN ACTION

Incandescent light bulbs

The complex relationship between electric current and temperature is put to use in a very common application: the incandescent light bulb (Figure 5.4.3).

FIGURE 5.4.3 An incandescent bulb produces light when its filament heats up.

An incandescent light bulb consists of a thin piece of curled or bent wire, called a *filament*, in a glass bulb. Often the bulb is evacuated (has the air removed) or filled with an inert (unreactive) gas so that the metal filament does not corrode. The wire is usually made of tungsten or another metal with a high melting point.

When an electric current passes through the filament, it heats up. This in turn increases the resistance of the filament, causing it to heat up further. The filament quickly becomes so hot that it starts to glow, radiating heat and light.

Traditionally, most household lighting was provided by incandescent light bulbs. However, this form of lighting is very inefficient. Only a small amount of the energy that goes into an incandescent light bulb is transformed into light: over 95% of the energy is lost as heat.

More recent inventions such as fluorescent tubes and LEDs (light-emitting diodes) are much more efficient and are increasingly being used as alternatives to incandescent light bulbs. Some examples of these are shown in Figure 5.4.4. In 2007, as a strategy to reduce carbon dioxide emissions, the Australian Government announced plans to phase out the use of incandescent bulbs.

FIGURE 5.4.4 Alternatives to incandescent light bulbs include fluorescent tubes, fluorescent bulbs and LED bulbs.

Many household electrical heating devices such as toasters and bar heaters work on a similar principle to the incandescent light bulb; although, in these situations, light is the unwanted or wasted energy.

OHM'S LAW

Georg Ohm (1789–1854) discovered that if the temperature of a metal wire was kept constant, the current flowing through it was directly proportional to the potential difference across it: mathematically, $I \propto V$. This relationship is known as Ohm's law. This relationship means that if the potential difference across a wire is doubled, for example, then the current flowing through the wire must also double. If the potential difference is tripled, then the current would also triple.

Ohm's law is usually written as:

$\Delta V = IR$

where ΔV is the potential difference in volts (V)

I is current in amps (A)

R is the constant of proportionality called resistance, in ohms (Ω).

This equation can be transposed to give a quantitative (mathematical) definition for resistance: $R = \frac{\Delta V}{I}$

If an identical potential difference produces two different sizes of current when separately connected to two light bulbs, then the resistance of the two light bulbs must differ. A higher current would mean a lower resistance of the light bulb, according to Ohm's law. This is because, when a conductor provides less resistance, more current can flow.

Worked example 5.4.1

USING OHM'S LAW TO CALCULATE RESISTANCE

When a potential difference of 3 V is applied across a piece of wire, 5 A of current flows through it. Calculate the resistance of the wire.	
Thinking	**Working**
Ohm's law is used to calculate resistance.	$\Delta V = IR$
Rearrange the equation to find R.	$R = \frac{\Delta V}{I}$
Substitute in the known values.	$R = \frac{3}{5}$ $= 0.6\,\Omega$

Worked example: Try yourself 5.4.1

USING OHM'S LAW TO CALCULATE RESISTANCE

An electric bar heater draws 10 A of current when connected to a 240 V power supply. Calculate the resistance of the element in the heater.

OHMIC AND NON-OHMIC CONDUCTORS

Conductors that obey Ohm's law are known as **ohmic** conductors. Ohmic conductors are usually called **resistors**.

An ohmic conductor can be identified by measuring the current that flows through the conductor when different potential differences are applied across it.

Worked example 5.4.2

USING OHM'S LAW TO CALCULATE RESISTANCE, CURRENT AND POTENTIAL DIFFERENCE

The table below shows measurements for the potential difference and corresponding current for an ohmic conductor.

ΔV [V]	0	2	4	ΔV_2
I [A]	0	0.25	I_1	0.75

Determine the missing results, I_1 and V_2.

Thinking	**Working**
Determine the factor by which potential difference has increased from the second column to the third column.	$\frac{4}{2} = 2$ The potential difference has doubled.
Apply the same factor increase to the current in the second column, to determine the current in the third column (I_1).	$I_1 = 2 \times 0.25$ $= 0.50\,\text{A}$
Determine the factor by which current has increased from the second column to the fourth column.	$\frac{0.75}{0.25} = 3$ The current has tripled.
Apply the same factor increase to the potential difference in the second column, to determine the potential difference in the fourth column (V_2).	$\Delta V_2 = 3 \times 2$ $= 6\,\text{V}$

Worked example: Try yourself 5.4.2

USING OHM'S LAW TO CALCULATE RESISTANCE, CURRENT AND POTENTIAL DIFFERENCE

The table below shows measurements for the potential difference and corresponding current for an ohmic conductor.

ΔV [V]	0	3	9	ΔV_2
I [A]	0	0.20	I_1	0.80

Determine the missing results, I_1 and V_2.

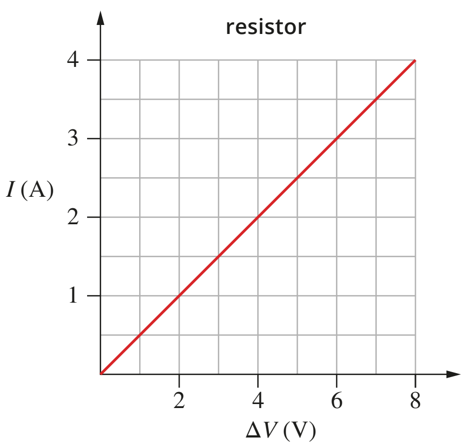


FIGURE 5.4.5 As the resistance of an ohmic conductor is constant, the I–ΔV graph is a straight line.

The data from an experiment in which the current and potential difference is measured for a device is usually plotted on an I–ΔV graph. If the conductor is ohmic, this graph will be a straight line, as can be seen in Figure 5.4.5.

The resistance of the ohmic conductor (or resistor) can be found from the gradient of the I–ΔV graph. Ohm recognised that the gradient was equal to the inverse of the resistance:

$$\frac{1}{R} = \frac{\text{rise}}{\text{run}} = \frac{4-1}{8-2} = \frac{3}{6}$$

$$\therefore R = \frac{6}{3} = 2\Omega$$

However, not all conductors are ohmic. The I–ΔV graphs for **non-ohmic** conductors are not straight lines (see Figure 5.4.6). Light bulbs and diodes are examples of non-ohmic conductors.

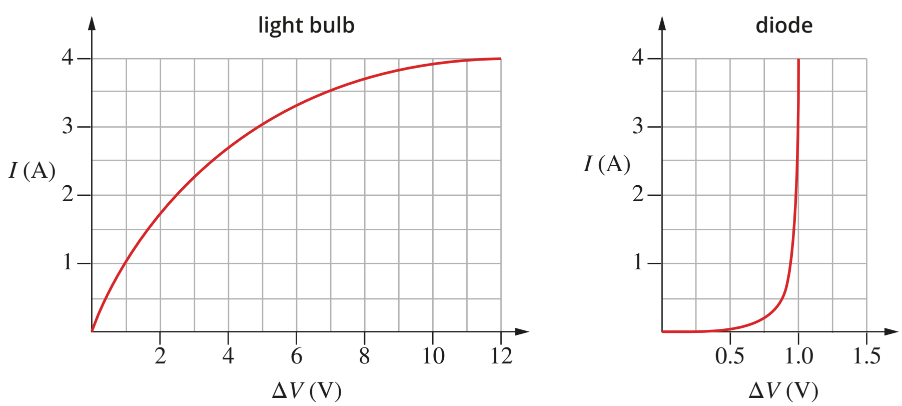


FIGURE 5.4.6 The I–ΔV graph for a non-ohmic resistor is not a straight line.

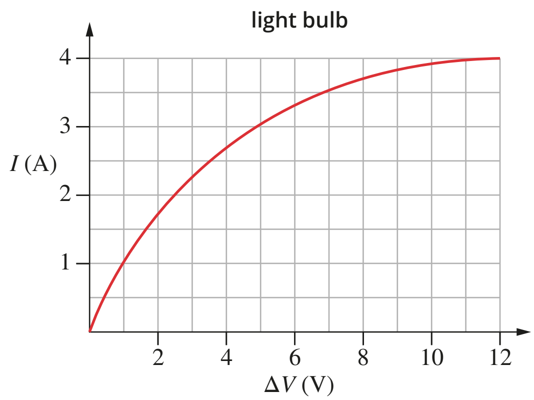
Using I–ΔV graphs to determine resistance

The inverse of resistance is defined as the ratio $I/\Delta V$. For an ohmic conductor, this value will be a constant regardless of the potential difference across the conductor. However, the resistance of a non-ohmic conductor will vary. The resistance of a non-ohmic conductor for a particular potential difference can be found by determining the current flowing through the conductor at this value.

Worked example 5.4.3

CALCULATING RESISTANCE FOR A NON-OHMIC CONDUCTOR

A light globe has the I–ΔV characteristics shown in the graph. Calculate the resistance of the light globe when the potential difference is 5.0 V.



Thinking	**Working**
From the graph, determine the current at the required potential difference.	At $\Delta V = 5\,V$, $I = 3\,A$
Substitute these values into Ohm's law to find the resistance.	$R = \frac{\Delta V}{I}$ $= \frac{5}{3}$ $= 1.67\,\Omega$

Worked example: Try yourself 5.4.3

CALCULATING RESISTANCE FOR A NON-OHMIC CONDUCTOR

A 240 V, 60 W incandescent light glove has the I–ΔV characteristics as shown in the graph. Calculate the resistance of the light globe when the potential difference is 175 V.

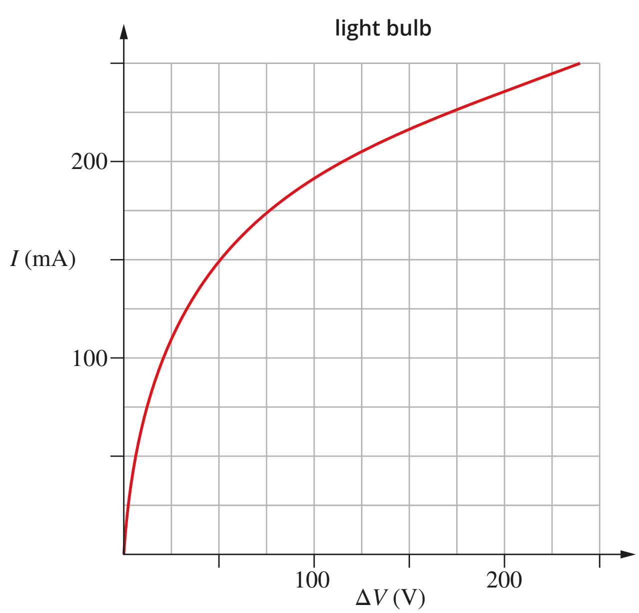


FIGURE 5.4.7 Common resistors are electrical devices with a known resistance. The coloured bands indicate the resistor's resistance and tolerance.

RESISTORS IN SIMPLE CIRCUITS

Ohmic resistors are often used to control the amount of current in a particular circuit. Resistors can be manufactured to produce a relatively constant resistance over a range of temperatures. A colour-coding system is used on resistors to explain the amount of resistance they provide, including a percentage tolerance (precision). Figure 5.4.7 shows a resistor that uses the colour-coding system.

PHYSICSFILE

Colour-coded resistors

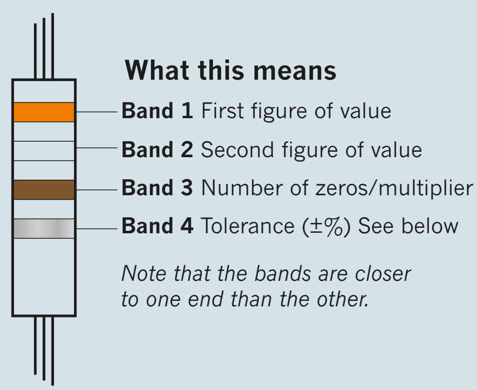
A resistor is typically a small piece of equipment which does not allow enough room to clearly print information about the resistor in the form of numbers. A colour-coding system is used on many resistors to convey detailed information in a small space about the resistance and tolerance of the resistor. Figure 5.4.8 below explains how to interpret the colour-coding system.

Resistor colour code

Band colour	Value
Black	0
Brown	1
Red	2
Orange	3
Yellow	4
Green	5
Blue	6
Purple	7
Grey	8
White	9
Gold	0.1
Silver	0.01

Tolerance colour code

Band colour	±%
Brown	1
Red	2
Gold	5
Silver	10
None	20



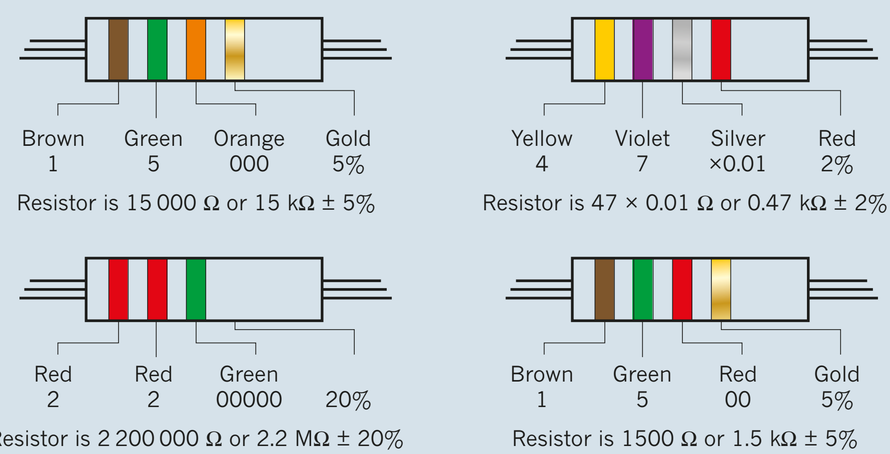


FIGURE 5.4.8 Examples of resistor colour-coding.

Ohm's law can be used to determine the current flowing through a resistor when a particular potential difference is applied across it. Similarly, if the current and resistance are known, the potential difference across the resistor can be calculated.

Worked example 5.4.4

USING OHM'S LAW TO FIND CURRENT

A 100 Ω resistor is connected to a 12 V battery. Calculate the current (in mA) that would flow through the resistor.

Thinking	Working
Recall Ohm's law.	$\Delta V = IR$
Rearrange the equation to make I the subject.	$I = \frac{\Delta V}{R}$
Substitute in the values for this problem and solve.	$I = \frac{12}{100}$ $= 0.12\ \text{A}$
Convert the answer to the required units.	$I = 0.12\ \text{A}$ $= 0.12 \times 10^3\ \text{mA}$ $= 120\ \text{mA}$

Worked example: Try yourself 5.4.4

USING OHM'S LAW TO FIND CURRENT

The element of a bar heater has a resistance of 25 Ω. Calculate the current (in mA) that will flow through this element if it is connected to a 240 V supply.

Worked example 5.4.5

USING OHM'S LAW TO FIND POTENTIAL DIFFERENCE

A current of 0.25 A flows through a 22 Ω resistor. Calculate the potential difference across the resistor. Give your answer correct to one decimal place.

Thinking	Working
Recall Ohm's law.	$\Delta V = IR$
Substitute in the known values and solve.	$\Delta V = 0.25 \times 22$ $= 5.5\ \text{V}$

Worked example: Try yourself 5.4.5

USING OHM'S LAW TO FIND POTENTIAL DIFFERENCE

The globe of a torch has a resistance of 5.7 Ω when it draws 700 mA of current. Calculate the potential difference across the globe.

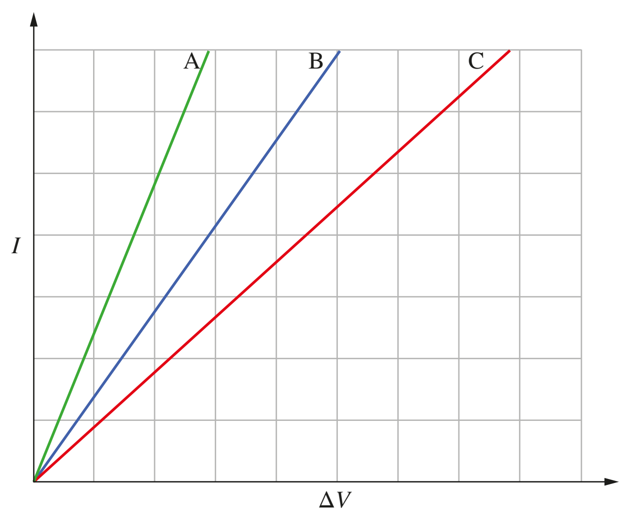
5.4 Review

SUMMARY

- Resistance is a measure of how hard it is for current to flow through a particular material. Resistance is measured in ohms (Ω).
- The resistance of a material depends on its length, cross-sectional area and temperature.
- Ohm's law describes the relationship between current, potential difference and resistance: $\Delta V = IR$
- Ohmic conductors have a constant resistance. The resistance of non-ohmic conductors varies for different potential differences.

KEY QUESTIONS

1 An experiment is conducted to gather data about the relationship between current and potential difference for three ohmic devices, labelled A, B and C. The data is used to plot an I–ΔV graph for each device, as shown below.



a For a given potential difference, list the devices in order of highest current to lowest current.

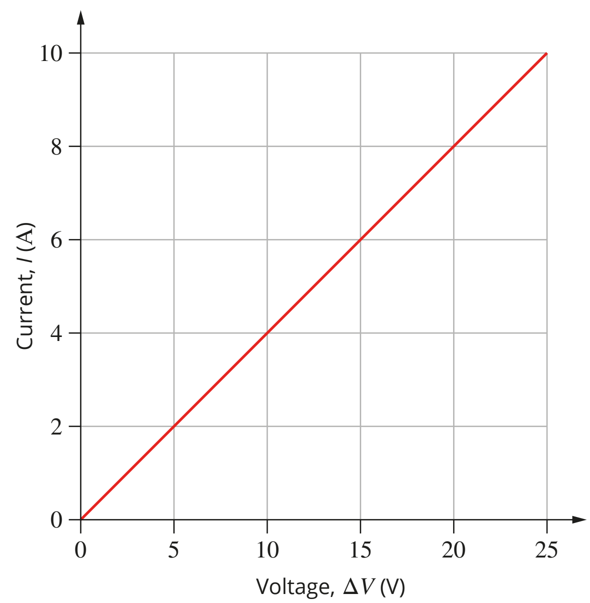
b List the devices in order of highest resistance to lowest resistance.

2 The table below shows measurements for the potential difference and corresponding current for an ohmic conductor.

ΔV [V]	0	2	3	ΔV_2
I [A]	0	0.25	I_1	0.60

Determine the missing results, I_1 and ΔV_2.

3 A student obtains the graph of the I–ΔV characteristics of a piece of resistance wire shown below.



a Explain whether this piece of wire is ohmic or non-ohmic.

b What current flows in this wire at a potential difference of 7.5 V?

c What is the resistance of this wire?

4 A student finds that the current through a resistor is 3.5 A when a potential difference of 2.5 V is applied to it.

a What is the resistance?

b The potential difference is then doubled and the current is found to increase to 7.0 A. Is the resistor ohmic or not?

5 Rose and Rachel are trying to find the resistance of an electrical device. They find that at 5 V it draws a current of 200 mA and at 10 V it draws a current of 500 mA. Rose says that the resistance is 25 Ω, but Rachel maintains that it is 20 Ω. Who is right and why?

6 Nick has an ohmic resistor to which he has applied 5 V. He measures the current as 45 mA. He then increases the potential difference to 8 V. What current will he find now?

7 Lisa finds that when she increases the potential difference across an ohmic resistor from 6 V to 10 V the current increases by 2 A.

 a What is the resistance of this resistor?

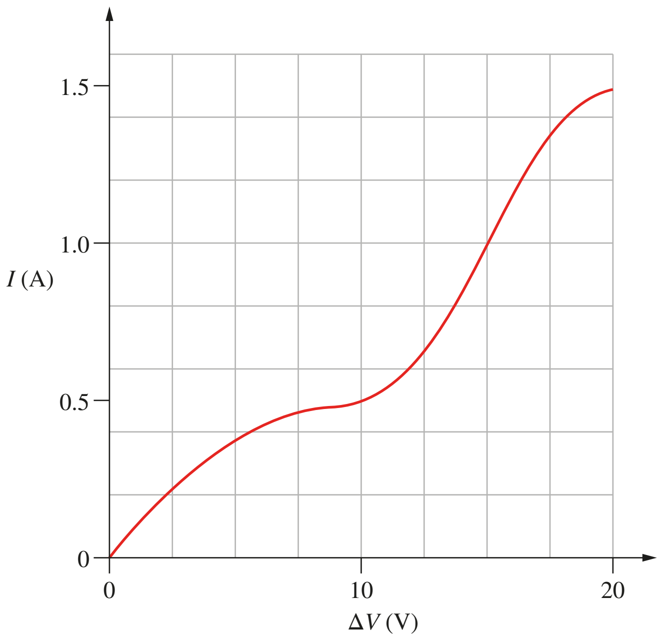
 b What current does it draw at 10 V?

8 The resistance of a piece of wire is found to be 0.8 Ω. What would be the resistance of:

 a a piece of the same wire twice as long?

 b a piece of wire of twice the diameter?

9 A strange electrical device has the I–ΔV characteristics shown in the graph below.



 a Is it an ohmic or non-ohmic device? Explain.

 b What current is drawn when a potential difference of 10 V is applied to it?

 c What potential difference would be required to double the current drawn at 10 V?

 d What is the resistance of the device at:

  i 10 V?

  ii 20 V?

5.5 Series and parallel circuits

Electric circuits are the basis of much of our modern society. This section introduces a range of circuits, from simple series circuits to the complex parallel wiring systems that make up a modern home. Electric circuits can be used to perform energy transfers and transformations through devices such as light bulbs, thermistors, light-dependent resistors and light-emitting diodes.

When a circuit contains more than one resistor, Ohm's law alone is not sufficient to predict the current flowing through and the potential difference across each resistor. Additional concepts such as Kirchhoff's rules and the idea of equivalent resistance can be used to analyse these complex, multi-component circuits.

No matter how complex a circuit, it can always be broken up into sections in which the circuit elements are in series or parallel. This section investigates the difference between these two types of circuits.

RESISTORS IN SERIES

Some circuits contain more than one electrical component. When these components are connected one after another in a continuous loop, this is called a **series circuit**. Components connected in this way are said to have been connected 'in series'. The circuit shown in Figure 5.5.1 shows a resistor and a light bulb connected in series with an electric cell.

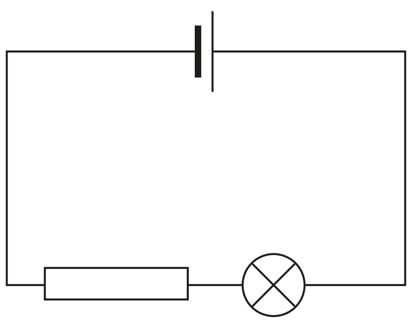


FIGURE 5.5.1 This circuit has a resistor and light bulb connected in series.

Series circuits are very easy to construct, but they have some disadvantages. As every component is connected one after the other, all components are dependent on each other. If one component is removed or breaks down, the circuit is no longer a closed loop and it won't work. This is referred to as an open circuit and the current stops. Figure 5.5.2 shows how removing a globe from a series circuit interrupts the entire circuit and breaks the current flow. Due to this characteristic, series circuits with more than one component are not commonly used in the home.

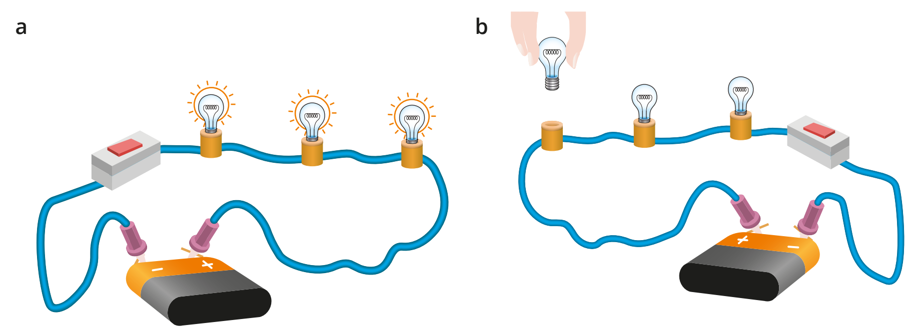


FIGURE 5.5.2 (a) There is no break in the circuit so the circuit is a closed loop. (b) When one light bulb is removed there is an open circuit.

Conservation of charge

When analysing a series circuit it is important to understand that the same amount of current flows in every part of the circuit. Since electric charges are not created or destroyed within an electric circuit, the current flowing out of the cell must be the same as the current flowing through the lamp, which is also the same as the current flowing through the resistor. This current also flows unchanged back into the cell as shown in Figure 5.5.3.

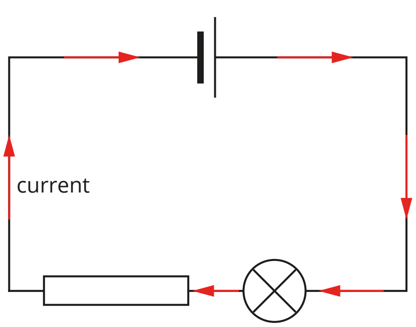


FIGURE 5.5.3 In a series circuit, the same current flows through each component.

> The current in a series circuit is the same in every part of the circuit.

Remember that, by convention, the current is represented as flowing from the positive terminal of the cell to the negative terminal. Electrons move in the opposite direction.

Kirchhoff's loop rule

Kirchhoff's loop rule says that the sum of the potential differences across all the elements around any circuit (loop) must be zero. This means that the total potential drop around a closed circuit must be equal to the total potential gain in the power source. For example, if a battery provides 9V to a circuit, then the sum of the potential drop across each of the components must add to 9V.

This rule is essentially another version of the law of conservation of energy.

> The energy given to the charges (potential gain) must be equal to the energy lost by the charges (potential drop). In a series circuit, the energy loss will be spread across the different components.

Figure 5.5.4 shows how the electrical potential provided by the battery is shared across a resistor and a lamp. The power supply in the figure is labelled EMF. Devices that are a source of energy for a circuit are referred to as sources of EMF or electromotive force. EMF, measured in volts (V), is another term for the work done on charges to provide a potential difference between the terminals of the power supply. In this series circuit, the sum of the potential drops across the resistor and the lamp (i.e. $V_1 + V_2$) will be equal to the potential difference (EMF) provided the resistance of the wires is negligible.

There are a number of ways to visualise the energy changes in this circuit. One common analogy is to think of the charges as water being pumped around an elevated water course. The water gains potential energy as it is pumped higher, and as it flows back down the potential energy is converted into other forms. The diagram in Figure 5.5.5 shows how the analogy works with the energy changes that occur in a circuit. The battery acts as a 'pump' that pushes electrons up to a higher energy level and they gain potential energy. As the electrons pass down through components in the circuit, their potential energy is transformed into other forms.

The change in electrical energy available to electrons can also be represented graphically, as shown in Figure 5.5.6. The electric potential energy available in the field changes as electrons move around the circuit. Some of this energy is lost as the electrons pass through the resistor. The remaining energy is lost as the electrons pass through the bulb. In this circuit, the bulb has more resistance than the resistor.

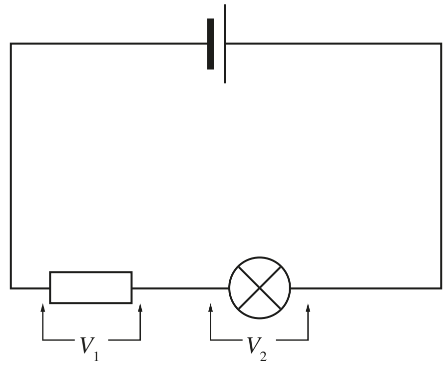


FIGURE 5.5.4 Kirchhoff's loop rule.

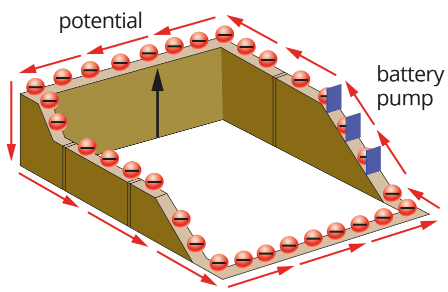


FIGURE 5.5.5 An analogy for analysing a circuit.: the battery acts as a 'pump' which transfers potential energy to electrons. The electrons lose potential energy as they flow 'down' through components in the circuit.

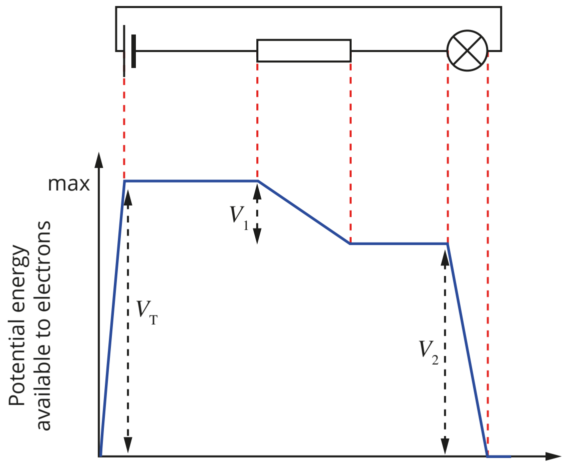


FIGURE 5.5.6 The electric potential energy of an electron changes as it moves around the circuit.

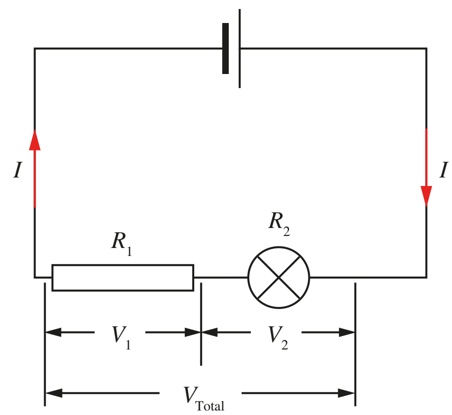


FIGURE 5.5.7 Ohm's Law can be used to show the total potential difference is the sum of the potential differences across the individual components.

Equivalent series resistance

Consider the circuit in Figure 5.5.7. If the resistance of the fixed resistor is R_1, the resistance of the lamp is R_2 and the current flowing through both of them is I, then Ohm's law gives:

$\Delta V_1 = IR_1$

and

$\Delta V_2 = IR_2$

The total potential difference drop across the two components is:

$\Delta V_{Total} = \Delta V_1 + \Delta V_2 = IR_1 + IR_2 = I \times (R_1 + R_2)$

This equation shows the relationship between the potential difference supplied by the cell and the potential differences of the lamp and resistor. The last part of the equation also shows that the lamp and resistor can be replaced with a single resistor, without changing the current in the circuit. The single resistor needs to have a total resistance of $R_1 + R_2$.

Using Ohm's law in such calculations, it is possible to show the relationship between the potential difference supplied by the cell, ΔV_{Total}, the current flowing in the circuit, I, and the resistances of the two components R_1 and R_2.

In general, a number of individual resistors connected in series can be replaced by an equivalent **effective resistance** (also called the total resistance, R_T) equal to the sum of the individual resistances. Figure 5.5.8 shows how two resistors can be replaced with a single one.

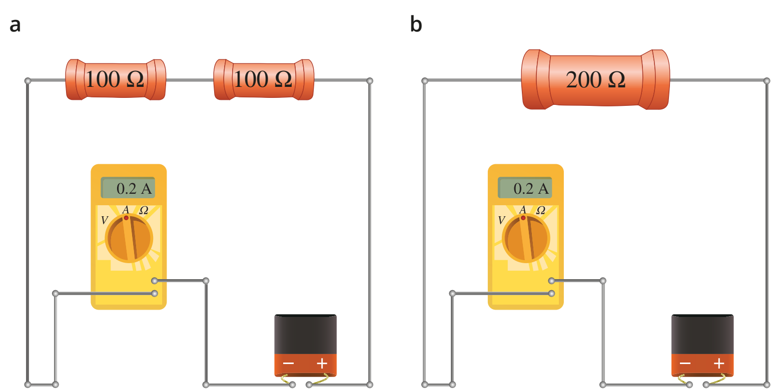


FIGURE 5.5.8 Two 100 Ω resistors (a) can be replaced with a single equivalent 200 Ω resistor (b) to have the same effect in a circuit.

Equivalent resistances can be used in circuit analysis to simplify a complicated circuit diagram so that current and potential difference can be determined.

$R_T = R_1 + R_2 + \ldots + R_n$

where R_T is the equivalent effective series resistance and R_1, R_2,…, R_n are the individual resistances.

Worked example 5.5.1

CALCULATING AN EQUIVALENT SERIES RESISTANCE

A 100 Ω resistor is connected in series with a 690 Ω resistor and a 1.2 kΩ resistor. Calculate the equivalent series resistance.

Thinking	Working
Recall the formula for equivalent series resistance.	$R_T = R_1 + R_2 + \ldots + R_n$
Substitute in the given values for resistance. Convert kΩ to Ω. Solve to find the equivalent series resistance.	$R_T = 100 + 690 + 1200 = 1990\,\Omega$

Worked example: Try yourself 5.5.1

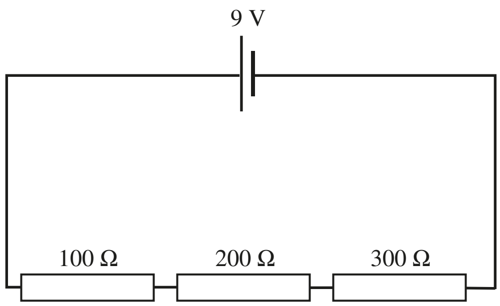
CALCULATING AN EQUIVALENT SERIES RESISTANCE

A string of Christmas lights consists of 20 light bulbs connected in series. Each bulb has a resistance of 8 Ω. Calculate the equivalent series resistance of the Christmas lights.

Worked example 5.5.2

USING EQUIVALENT SERIES RESISTANCE FOR CIRCUIT ANALYSIS

Use an equivalent series resistance to calculate the current flowing in the series circuit below and the potential difference across each resistor.

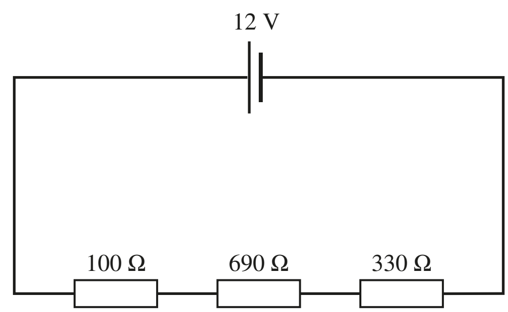


Thinking	Working
Recall the formula for equivalent series resistance.	$R_T = R_1 + R_2 + R_3 + \ldots + R_n$
Find the equivalent (total) resistance in the circuit.	$R_T = 100 + 200 + 300 = 600\,\Omega$
Use Ohm's law to calculate the current in the circuit. Whenever calculating current in a series circuit, use R_T and the potential difference of the power supply.	$I = \frac{\Delta V}{R}$ $= \frac{9}{600}$ $= 0.015\,\text{A}$
Use Ohm's law to calculate the potential difference across each separate resistor.	$\Delta V = IR$ Therefore $\Delta V_1 = 0.015\,\text{A} \times 100\,\Omega = 1.5\,\text{V}$ $\Delta V_2 = 0.015\,\text{A} \times 200\,\Omega = 3.0\,\text{V}$ $\Delta V_3 = 0.015\,\text{A} \times 300\,\Omega = 4.5\,\text{V}$
Use the loop rule to check the answer.	$\Delta V_T = \Delta V_1 + \Delta V_2 + \Delta V_3$ $= 1.5\,\text{V} + 3.0\,\text{V} + 4.5\,\text{V}$ $= 9.0\,\text{V}$ Since this is the same as the potential difference provided by the cell, the answer is reasonable.

Worked example: Try yourself 5.5.2

USING EQUIVALENT SERIES RESISTANCE FOR CIRCUIT ANALYSIS

Use an equivalent series resistance to calculate the current flowing in the series circuit below and the potential difference across each resistor.



RESISTORS IN PARALLEL

One of the disadvantages of series circuits is that if a switch is opened or a device disconnected, then the circuit is broken and current stops flowing. In everyday life, we often want to switch devices on and off independently. **Parallel circuits** allow us to do this.

The circuit diagram in Figure 5.5.9 shows a simple parallel circuit. Even if switch A is open (as shown), lamp B will still light up as it is part of an unbroken complete circuit including the battery. Similarly, if switch A is closed and switch B is opened, current will light up lamp A and not lamp B. Alternatively, both switches could be closed to light up both lamps or both switches could be opened to switch both lamps off.

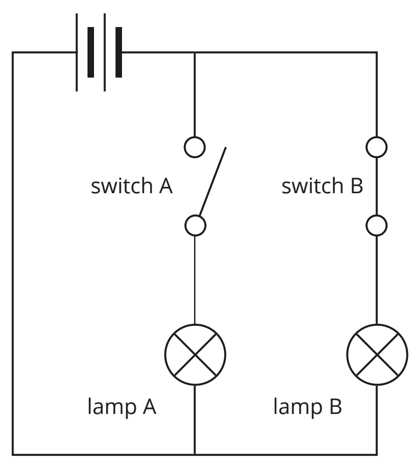


FIGURE 5.5.9 In this parallel circuit, lamp A will be off and lamp B will be on.

In a series circuit, all the components are in the same loop and therefore the same current flows through each component. In comparison, each loop of a parallel circuit acts like an independent circuit with its own current.

Consider again the shopping mall analogy. When a certain number of people are on the first floor they can get to the ground floor by either going down the lift or walking down the stairs. One way is easier than the other, so this represents a difference in resistance between two parallel paths. Most people will take the path of least resistance (the lift) and fewer people will take the path of greater resistance (the stairs). The rate at which people enter the lift plus the rate at which they enter the stairs combines to equal the rate at which people leave the first floor. When the people re-join, the rate at which people leave the lift and stairs combine to be the rate at which people enter the ground floor. This is the same rate as it was for people leaving the first floor.

The same occurs with charges flowing in a parallel circuit. Figure 5.5.10 shows how the charges go through one globe or the other. This means that while the current before and after the components remains constant, in the parallel section, the current is divided between each branch. The readings on both ammeters, A_1 and A_2, will be the same.

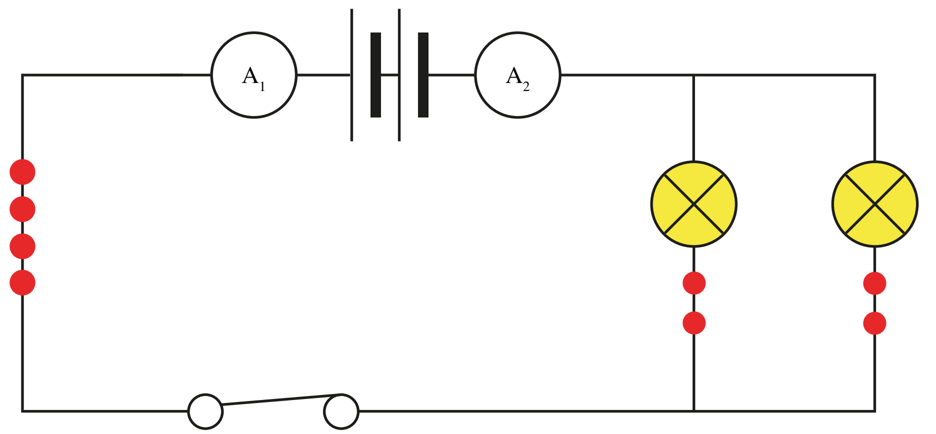


FIGURE 5.5.10 Charges flowing in a parallel circuit will flow down one of two paths.

Unlike a series circuit, in a parallel circuit the potential difference is not shared between resistors; the potential difference is the same across each branch. This is because, while the charges take different pathways, they have the same change in energy no matter which path they take. The electrical potential in the wire, before the parallel circuit components, is the same no matter the path and the electrical potential in the wire after the parallel components is also the same, so the difference in electrical potential must also be the same no matter which path the current takes. An advantage of parallel circuits is that globes connected in this way are brighter than if they were connected in series. The potential energy of the charges is not shared between the globes.

> The current in the main part of a parallel circuit is the sum of the currents in each branch of the circuit.
>
> $I_T = I_1 + I_2 + ... + I_n$

Using the shopping mall analogy, the people start on the first floor and walk along until some walk towards the lift and some walk towards the stairs, with both groups are still on the first floor. In the circuit, electrons in the experience the same electrical potential whether they flow along one parallel branch or the other. As the people get to the ground floor, via the lift or the stairs, they lose energy from the field (in this case, gravitational potential energy). In the circuit, energy is transferred from the electric field via the electrons into other forms of energy. When the people reach the ground floor via either the lift of the stairs, they are once again at the same level as each other and they can re-join each other and continue along the mall. In the electric circuit, the electrons passing from the resistors are at the same potential as each other (but are at a lower potential than the electrons before the resistors) and they can continue to move along the circuit.

> The potential difference is the same in each branch of a parallel circuit.

Kirchhoff's junction rule

Parallel circuits involve **junctions** where current can flow in a variety of directions. The behaviour of current at these points is predicted by Kirchhoff's junction rule:

> The total amount of current flowing into a junction must be the same as the total current flowing out of the junction.

This rule is just an extension of the idea of conservation of charge; that is, that charges cannot be created or destroyed. Although the number of electrons flowing into a junction might be very large, electrons are not created or destroyed in the junction so the same number of electrons must flow out again. This is illustrated in Figure 5.5.11. Kirchhoff's junction rule explains how current splits in a parallel circuit. It explains why the current in the main part of the circuit is the sum of the currents in each parallel branch.

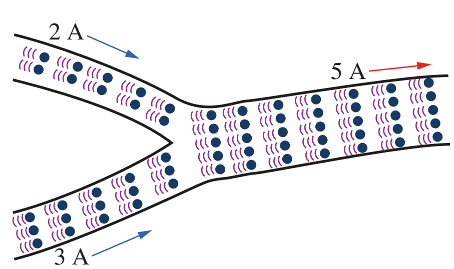


FIGURE 5.5.11 The current flowing into any junction must be equal to the current flowing out of it.

Equivalent parallel resistance

When additional resistors are added into a series circuit the total resistance of the circuit increases. As the potential difference across the combination of resistors remains the same, increased resistance means that less current flows through the circuit.

In contrast, adding an additional resistor in parallel means that more current flows through the circuit because another path for the charges to take has been added. The potential difference across each resistor in the parallel combination is the same so this means that the total resistance for the circuit decreases such that it is always less than the value of the least resistive branch.

Consider the shopping mall analogy once again. If there were only two ways to get to the ground floor from the first floor, then there would be a limit to the rate at which people could get to the ground level. However, if another lift were added, then more people could get to the ground in the same period of time, so the rate at which people could get to the ground floor would be greater. There would be less 'resistance' for people trying to get to the ground floor.

To represent the addition of resistors in a mathematical way, such that the overall resistance decreases, you can use the following relationship:

$$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}$$

where R_T is the equivalent effective resistance and $R_1, R_2, \ldots, R_n$ are the individual resistances.

Worked example 5.5.3

CALCULATING AN EQUIVALENT PARALLEL RESISTANCE

A 100 Ω resistor is connected in parallel with a 300 Ω resistor. Calculate the equivalent parallel resistance.	
Thinking	**Working**
Recall the formula for equivalent effective resistance.	$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}$
Substitute in the given values for resistance.	$\frac{1}{R_T} = \frac{1}{100} + \frac{1}{300}$
Solve for R_T.	$\frac{1}{R_T} = \frac{1}{100} + \frac{1}{300}$ $= \frac{3}{300} + \frac{1}{300}$ $= \frac{4}{300}$ $R_T = \frac{300}{4}$ $= 75\,\Omega$

Worked example: Try yourself 5.5.3

CALCULATING AN EQUIVALENT PARALLEL RESISTANCE

A 20 Ω resistor is connected in parallel with a 50 Ω resistor. Calculate the equivalent parallel resistance.

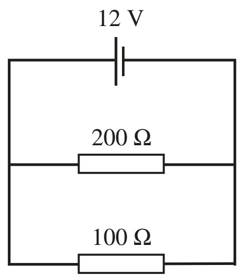
Notice that in the previous worked example, the equivalent effective resistance was smaller than the smallest individual resistance. This is because adding a resistor provides an additional pathway for current. Since more current flows, the resistance of the circuit has been effectively reduced.

If you consider the smallest resistor in any parallel combination, say, the 20 Ω resistor in Worked example: Try yourself 5.5.3, the addition of the 50 Ω resistor in parallel with it allows the current an extra pathway and therefore it is easier for the current to flow through the combination. The effective resistance of the pair must be less than the 20 Ω alone.

Worked example 5.5.4

USING EQUIVALENT PARALLEL RESISTANCE FOR CIRCUIT ANALYSIS

Find the equivalent parallel resistance to calculate the current flowing out of the 12 V cell in the parallel circuit shown. Then find the current flowing through each resistor.



Thinking	Working
Recall the formula for equivalent parallel resistance.	$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}$
Substitute in the given values for resistance.	$\frac{1}{R_T} = \frac{1}{100} + \frac{1}{200}$
Solve for R_T.	$\frac{1}{R_T} = \frac{1}{100} + \frac{1}{200}$ $= \frac{2}{200} + \frac{1}{200}$ $= \frac{3}{200}$ $R_T = \frac{200}{3}$ $= 67\,\Omega$
Use Ohm's law to calculate the current in the circuit. To calculate I, use the potential difference of the power supply and the total resistance.	$I_{circuit} = \frac{\Delta V}{R} = \frac{12}{67} = 0.18$ A
Use Ohm's law to calculate the current through each separate resistor. Remember that the potential difference across each resistor is the same as the potential difference of the power supply, 12 V in this case.	100 Ω resistor: $I_{100} = \frac{\Delta V}{R} = \frac{12}{100} = 0.12$ A 200 Ω resistor: $I_{200} = \frac{\Delta V}{R} = \frac{12}{200} = 0.060$ A
Use the junction rule to check the answers.	$I_{circuit} = I_{100} + I_{200}$ 0.18 A = 0.12 A + 0.060 A This is correct, so the answers are reasonable.

The effective (total) resistance of a set of resistors connected in parallel will always be smaller than the smallest resistor in the set.

In a parallel circuit:

$R_{Total} < R_{smallest\ resistor}$

PHYSICSFILE

Kirchhoff's contributions

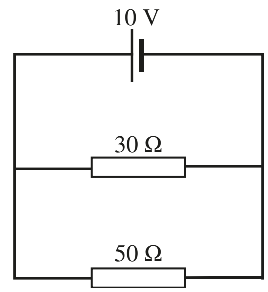
Both the junction rule discussed here and the loop rule described earlier were first discovered by the German physicist Gustav Kirchhoff (1824–87). Kirchhoff also made important contributions in the fields of spectroscopy, thermochemistry and the study of black-body radiation. He worked with Robert Bunsen, the German chemist who developed the Bunsen burner.

FIGURE 5.5.12 Gustav Kirchhoff discovered the rules that underpin our understanding of how electric circuits work.

Worked example: Try yourself 5.5.4

USING EQUIVALENT PARALLEL RESISTANCE FOR CIRCUIT ANALYSIS

Find the equivalent parallel resistance to calculate the current flowing out of the 10 V cell in the parallel circuit shown. Then find the current flowing through each resistor.



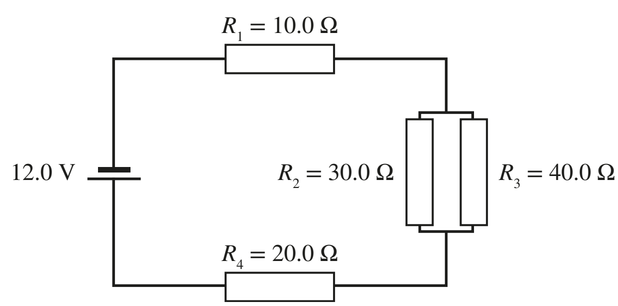
COMPLEX CIRCUIT ANALYSIS

Some circuits combine elements of series wiring and parallel wiring. A general strategy for analysing these circuits is to reduce the complex circuit to a single equivalent resistance to determine the current drawn by the circuit. It is then possible to step back through the process of simplification to analyse each section of the circuit as needed.

Worked example 5.5.5

COMPLEX CIRCUIT ANALYSIS

Calculate the potential difference across and current through each resistor in the circuit below.



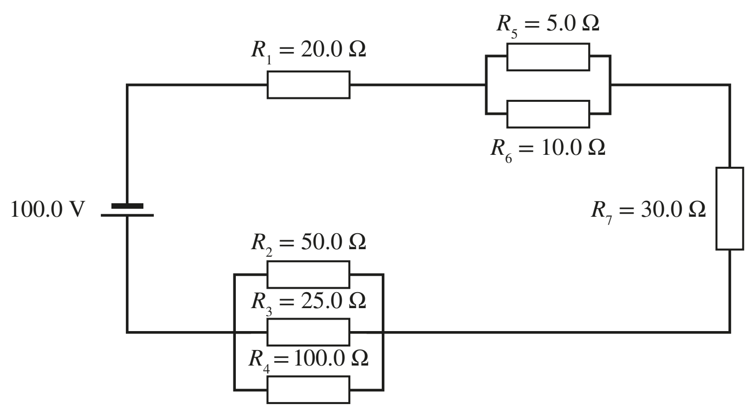
Thinking	**Working**
Find an equivalent resistance for the parallel resistors. The effective resistance of these should be less than the smaller resistor, that is, less than 30 Ω.	$\frac{1}{R_{2-3}} = \frac{1}{R_2} + \frac{1}{R_3}$ $= \frac{1}{30.0} + \frac{1}{40.0}$ $= \frac{4}{120.0} + \frac{3}{120.0}$ $R_{2-3} = \frac{120.0}{7} = 17.1\,\Omega$
Find an equivalent series resistance for the circuit as the circuit can now be thought of as three resistors in series: 10.0 Ω, 17.1 Ω and 20.0 Ω.	$R_T = 10.0\,\Omega + 17.1\,\Omega + 20.0\,\Omega$ $= 47.1\,\Omega$

Use Ohm's law to calculate the current in the circuit. Use the potential difference of the power supply and the total resistance to do this calculation.	$\Delta V = IR$ $I = \frac{\Delta V}{R}$ $= \frac{12.0}{47.1}$ $= 0.255\,A$
Use Ohm's law to calculate the potential difference across each resistor (or parallel group of resistors) in series. (Note that the potential difference across R_2 is the same as that across R_3 as they are in parallel.)	$\Delta V = IR$ $\Delta V_1 = 0.255 \times 10.0 = 2.55\,V$ $\Delta V_{2-3} = 0.255 \times 17.1 = 4.36\,V$ $\Delta V_4 = 0.255 \times 20.0 = 5.10\,V$ Check: $2.55 + 4.36 + 5.10 \approx 12.0\,V$ (with some slight rounding error) This confirms that the loop rule holds for this circuit.
Use Ohm's law where necessary to calculate the current through each resistor.	$I_1 = I_4 = 0.255\,A$ $I = \frac{\Delta V}{R}$ $I_2 = \frac{4.36}{30.0} = 0.145\ A$ $I_3 = \frac{4.36}{40.0} = 0.109\ A$ Check: $0.145 + 0.109 \approx 0.255\,A$ (with some slight rounding error) This confirms that the junction rule holds for this section.

Worked example: Try yourself 5.5.5

COMPLEX CIRCUIT ANALYSIS

Calculate the potential difference across and the current through each resistor in the circuit below.



PHYSICSFILE

Identical resistors in parallel

Where identical resistors are placed in parallel, the total resistance of the combination can be found by simply dividing the value of one of the resistors by the number of resistors.

For example, the total resistance of three $12\,\Omega$ resistors connected in parallel would have an effective resistance of $4\,\Omega$.

$R_T = 12 \div 3 = 4\,\Omega$.

The three $12\,\Omega$ resistors in parallel could be replaced with a single $4\,\Omega$ resistor.

Similarly, the equivalent resistance of two $10\,\Omega$ resistors placed in parallel would be $5\ \Omega$.

PHYSICS IN ACTION

Superconductors

In 1908 Dutch physicist Kamerlingh Onnes (1853–1926) was the first to liquefy helium (Figure 5.5.13). This occurred at 4.2 K (–269°C). Onnes was the first to liquefy a number of gases. He was also the first person to achieve the then world-record lowest temperature of 1.5 K.

Onnes began investigating the resistance of pure metals at very low temperatures. Some scientists believed at the time that the resistance of metals would greatly increase, or even become infinite near absolute zero (–273°C). Other scientists, Onnes among them, believed that the electrical resistance would eventually drop to nil.

In 1911, Onnes immersed a solid wire of mercury into liquid helium and found that at 4.2 K its resistance was indeed zero. He called this the *superconducting state*, a new state of matter. Theoretically, once an electric current was started in a loop maintained at superconducting temperatures, it would circulate indefinitely. Onnes was awarded the Nobel Prize in Physics in 1913.

Other metals were soon found to become superconductors at extremely low temperatures—for example, aluminium at 1.2 K and lead at 7.9 K. However, not all metals can become superconducting. Perhaps surprisingly, the excellent electrical conductor copper does not become superconducting at any temperature.

The mechanism by which a metal's electrical resistance drops to zero was not understood at the time. In 1957 three American physicists, John Bardeen, Leon Cooper and Robert Schrieffer, published a paper that explained the phenomenon. The explanation is now known as the BCS theory and required quantum mechanics, not known in Onnes' era. Bardeen, Cooper and Schrieffer were awarded the 1972 Nobel Prize in Physics.

Superconductivity hit the headlines again in the late 1980s when some ceramic compounds were discovered to become superconducting at relatively high temperatures. Despite the fact that copper by itself is not a superconductor, most high-temperature superconductors (HTS) are compounds of copper. An example is the compound yttrium barium copper oxide, $YBa_2Cu_3O_7$, which becomes superconducting at liquid nitrogen temperatures (77 K or –196°C). This new class of ceramic superconductors generated a great deal of interest, and currently the record stands at 135 K (–138°C) using the compound mercury barium calcium copper oxide, $HgBa_2Ca_2Cu_3O_8$.

The mechanism by which this class of ceramic materials becomes superconducting is not well understood and so they are currently a very active area of both theoretical and experimental research.

The eventual goal is to discover a material that will become superconducting at room temperature. Such a material would have many applications and would not require all the equipment and expense to maintain extremely low or high temperatures.

Superconductors have made it possible to produce extremely powerful magnets. Such magnets find applications in magnetic resonance imaging (MRI) machines used in medical scanning, in maglev (magnetic levitation) trains that float above the rails and therefore do not experience friction and clunking from the rails, and in the Large Hadron Collider to bend extremely high speed protons around the beam-line.

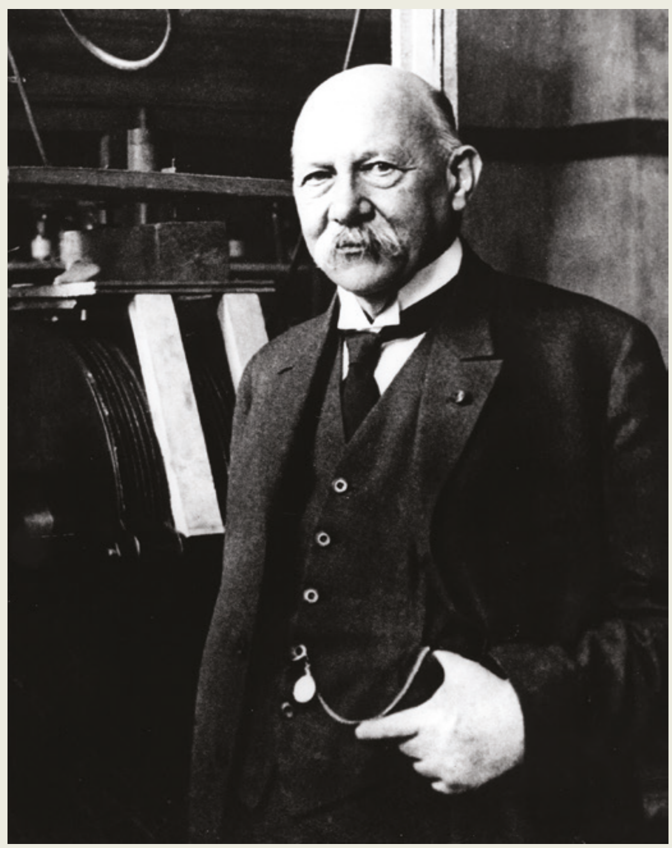


FIGURE 5.5.13 The Dutch physicist Kamerlingh Onnes.

RESISTORS AND POWER

A particular combination of resistors will draw different amounts of power depending on whether the resistors are wired in series or parallel. In general, since resistors in parallel circuits will draw more current than resistors in series circuits, parallel circuits use more power than series circuits containing the same resistors.

Recall from Section 5.3 that the equation for power is:

$P = \Delta V \times I$

where P is the power (W)

ΔV is the potential difference (V)

I is the current (A).

Worked example 5.5.6

COMPARING POWER IN SERIES AND PARALLEL CIRCUITS

Consider a 100 Ω and a 300 Ω resistor wired in parallel with a 12 V cell. Calculate the power drawn by these resistors. Compare this to the power drawn by the same two resistors when wired in series.	
Thinking	**Working**
Calculate the equivalent resistance for the parallel circuit.	$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}$ $= \frac{1}{100} + \frac{1}{300}$ $= \frac{3}{300} + \frac{1}{300}$ $= \frac{4}{300}$ $R_T = \frac{300}{4}$ $= 75\,\Omega$
Calculate the total current drawn by the parallel circuit.	$\Delta V = IR$ $I = \frac{\Delta V}{R} = \frac{12}{75} = 0.16$ A
Use the power equation to calculate the power drawn by the parallel circuit.	$P = \Delta VI$ $= 12 \times 0.16$ $= 1.92$ W
Calculate the equivalent resistance for the series circuit.	$R_T = R_1 + R_2 + \ldots + R_n$ $= 100 + 300$ $= 400\,\Omega$
Calculate the total current drawn by the series circuit.	$\Delta V = IR$ $I = \frac{\Delta V}{R} = \frac{12}{400} = 0.03$ A
Use the power equation to calculate the power drawn by the series circuit.	$P = \Delta VI$ $= 12 \times 0.03$ $= 0.36$ W
Compare the power drawn by the two circuits.	$\frac{P_{parallel}}{P_{series}} = \frac{1.92}{0.36} = 5.33$ The parallel circuit draws over 5 times as much power as the series circuit.

Worked example: Try yourself 5.5.6

COMPARING POWER IN SERIES AND PARALLEL CIRCUITS

Consider a 200 Ω and an 800 Ω resistor wired in parallel with a 12 V cell. Calculate the power drawn by these resistors. Compare this to the power drawn by the same two resistors when wired in series.

In Worked example 5.5.6 and Worked example: Try yourself 5.5.6, Ohm's law and the power equation were used separately to determine the power of a circuit component. You can combine the two equations to derive two useful equations:

$P = \Delta VI$ and $\Delta V = IR$

so

$P = (IR) \times I$

$P = I^2R$

Similarly

$P = \Delta VI \quad$ and $\quad I = \dfrac{\Delta V}{R}$

so

$$P = \Delta V\left(\frac{\Delta V}{R}\right)$$

$$P = \frac{\Delta V^2}{R}$$

PHYSICS IN ACTION

High power–low power

Simple heaters of various sorts often have a 'three heat' switch. An electric blanket will usually have 'low', 'medium' and 'high' settings, for example. Rather than making three different heating elements, the manufacturer can use two elements in different series and parallel combinations to obtain the three heat settings. If the two elements are placed in series the total resistance is relatively high and therefore the power will be a minimum, as $P = \dfrac{\Delta V^2}{R}$. For the medium setting one of the elements will be used by itself. The high setting is then achieved by placing both elements in parallel.

It is a simple matter to work out the relative power being used for the three settings. If it is assumed that the resistance of both elements is the same (R) and does not change appreciably with temperature, the effective resistance in the three cases will be given by:

Low heat (two elements in series): $R_T = R + R = 2R$

Medium heat (one element only): $R_T = R$

High heat (both in parallel): $R_T = \dfrac{1}{2}R$.

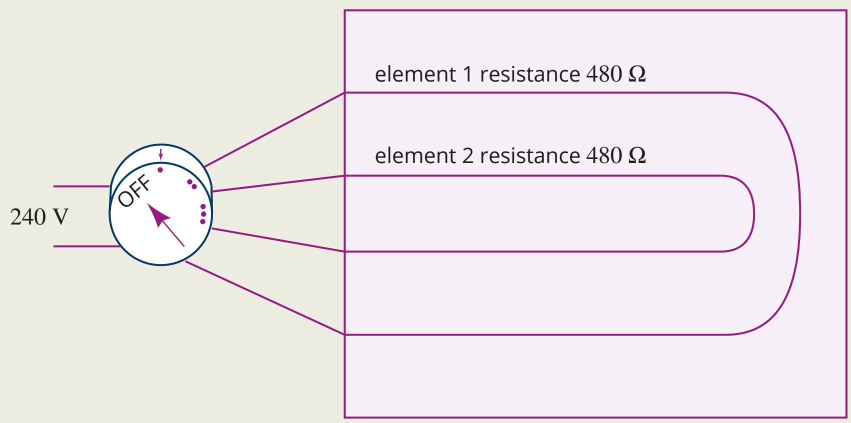


FIGURE 5.5.14 An example of the use of series and parallel combinations of resistors to achieve three heat settings for an electric blanket.

As the power is inversely proportional to the resistance $\left(P = \dfrac{\Delta V^2}{R}\right)$, if we call the high setting 100%, then the other two will be 50% and 25%.

To obtain the various settings the following circuits are used:

OFF Not connected to power

- Resistors connected in series, $R = 960\,\Omega$, $I = 0.25$ A, $P = 60$ W
- Only one resistive element is connected, $R = 480\,\Omega$, $I = 0.5$ A, $P = 120$ W
- Resistors connected in parallel, $R = 240\,\Omega$, $I = 1.0$ A, $P = 240$ W

PHYSICS IN ACTION

Parallel connections in the home

All household appliances and lights are connected in parallel. This is done for two reasons.

Figure 5.5.15 shows a TV, air-conditioner, heater and washing machine connected in series. Each of these devices is designed to operate at 240 V.

A circuit designed like the one in Figure 5.5.15 poses many problems. Firstly, all of the devices need to be switched on for the circuit to operate.

Secondly, the 240 V supplied to the circuit needs to be shared among all the components. Each component in the circuit would receive far less than the 240 V they require to operate. Also, as more and more devices are added to the circuit, the share of the 240 V would become smaller. This system could never be practical.

The circuit diagram in Figure 5.5.16 shows how the same devices could be connected in parallel.

Each device in the parallel circuit receives the same potential difference, 240 V. Each device can be independently switched on or off without affecting the others and more devices can be added to this system without affecting the operation of the others. The only practical issue with devices in a parallel circuit such as this is the total current they draw. This cannot exceed the capacity of the protection device in the circuit. This aspect of electric circuits will be dealt with in detail in the next section of this book.

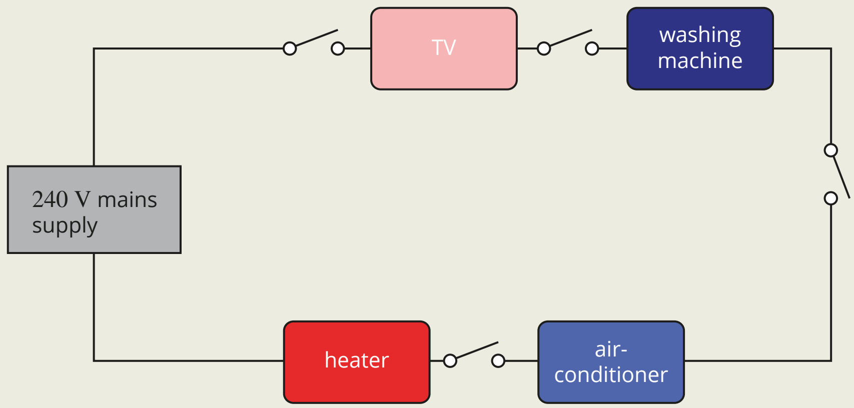


FIGURE 5.5.15 Household appliances connected in series.

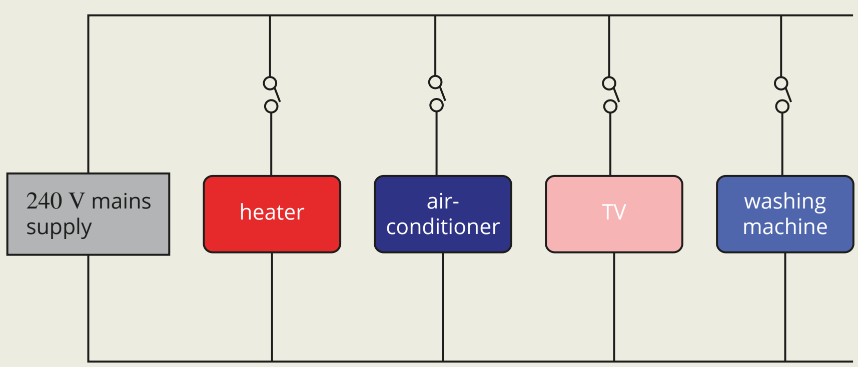


FIGURE 5.5.16 Household appliances connected in parallel.

5.5 Review

SUMMARY

- When resistors are connected in series, the:
  - current through each resistor is the same
  - sum of the potential differences is equal to the potential difference provided to the circuit
  - equivalent effective resistance R_T is equal to the sum of the individual resistances.
- Parallel circuits allow individual components to be switched on and off independently.
- When resistors are connected in parallel, the:
  - potential difference across each resistor is the same
  - current is shared between the resistors
  - equivalent effective resistance is given by the equation:

$$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots + \frac{1}{R_n}$$

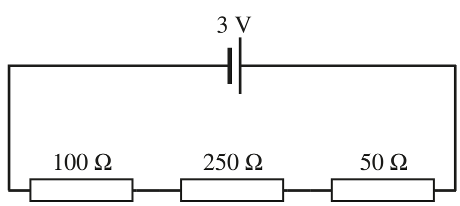
- Complex circuit analysis may require the calculation of both equivalent series and equivalent parallel resistances.
- A parallel circuit generally draws more power than a series circuit using the same resistors.

KEY QUESTIONS

1 Two 20 Ω resistors are connected in series with a 6 V battery. What is the potential difference across each resistor?

A 0.3 V
B 3 V
C 6 V
D 12 V

2 Consider the series circuit below.



a Calculate the current flowing in the circuit. Give your answer correct to two significant figures.
b Calculate the potential difference across the 100 Ω resistor in the circuit.

3 Two equal resistors are connected in parallel and are found to have an equivalent resistance of 68 Ω. Calculate the resistance of each resistor.

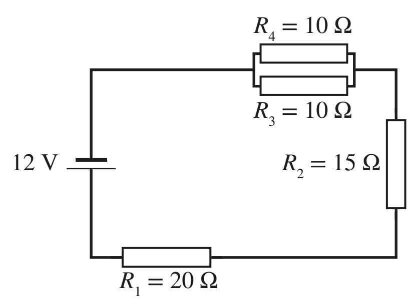
4 A 20 Ω resistor and a 10 Ω resistor are connected in parallel to a 5 V battery. Give your answers correct to two decimal places.

a Calculate the current drawn from the battery.
b Calculate the current flowing through the 20 Ω resistor.
c Calculate the current flowing through the 10 Ω resistor.

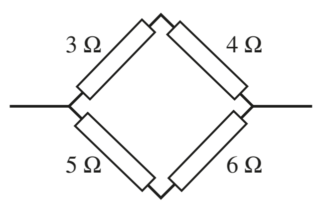
5 A 40 Ω resistor and a 60 Ω resistor are connected in parallel to a battery, with 300 mA flowing through the 40 Ω resistor.

a Calculate the potential difference of the battery.
b Calculate the current flowing through the 60 Ω resistor.

6 Calculate the potential difference across, and the current through, each resistor in the circuit below.



7 Calculate the equivalent resistance of the combination of resistors shown below.



8 Four 20 Ω light bulbs are connected to a 10 V battery. What is the total power output of the circuit if the light bulbs are connected:

a in series?
b in parallel?

9 Why are household circuits wired in parallel?

A to reduce the amount of expensive copper wire used
B to reduce the amount of current drawn by the household
C to allow appliances to be switched on and off independently
D to reduce the amount of electrical energy used by the household

5.6 Electrical safety

Homes, schools and workplaces are filled with all sorts of electrical appliances. You use these appliances every day. A scientific understanding of electricity can help you to understand how they work and how to make sure you use them safely and effectively.

ELECTRICITY IN THE HOME

Circuits in the home

The wiring for a house is much more complicated than the relatively simple series and parallel circuits considered so far. Figure 5.6.1 shows the basic structure of the electrical wiring in a house. Most appliances and power points are wired in parallel to allow them to be switched on and off independently of each other.

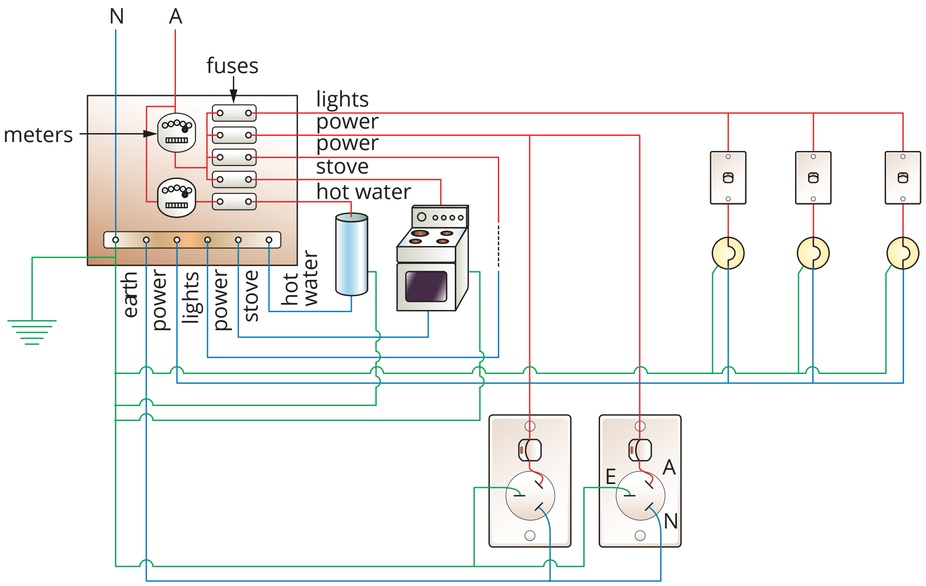


FIGURE 5.6.1 A household wiring diagram includes active (A), neutral (N) and earth (E) wires.

Power points in Australia have three pins. Each of these pins is connected to a different wire. Figure 5.6.1 shows the arrangement of the active, neutral and **earth** pins on a typical power point. The wires carrying the electric current to and from the appliance are known as the active wire (usually red or brown) and the neutral wire (usually black or blue). The third wire is an important safety feature called the earth wire (usually green or green and yellow). Figure 5.6.2 shows the corresponding active, neutral and earth pins in a power cord. This wiring pattern is an Australian Standard.

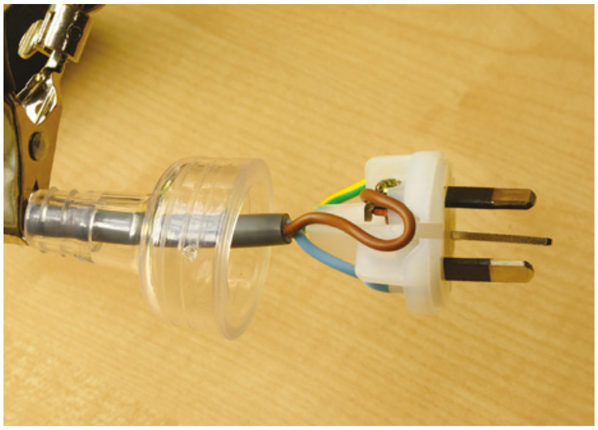


FIGURE 5.6.2 This three-pin plug shows the correct colour code and pin combination used.

The purpose of household electrical circuits is to enable electrical energy to be transferred to electrical appliances, where it is transformed into a range of other useful forms of energy. For example, an electric oven converts electrical energy into heat, whereas fans convert electrical energy into kinetic energy. Power points give users the option of connecting their own appliances and therefore choosing the type of energy produced.

Figure 5.6.3 shows that, before being distributed to various parts of the house, the active wires pass through meters that measure the amount of electrical energy supplied to the household.

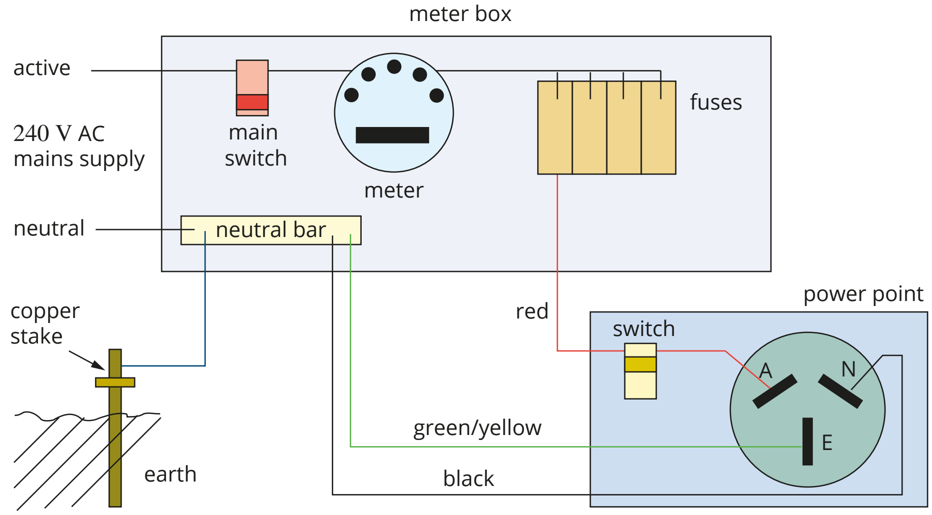


FIGURE 5.6.3 The active wire passes through the meter box before being connected to power points throughout the house. The circuit is completed by the neutral wire, returning through the neutral bar.

AC/DC

The electrical current that comes out of a power point is very different to the current that comes from a battery or electric cell (Figure 5.6.4). A power point provides alternating current (AC) whereas a car battery provides direct current (DC).

FIGURE 5.6.4 A power point and a battery.

For a start, the average potential difference between the active and neutral pins of a household power point is 240V. This is much higher than most electric cells or batteries can provide. In fact, 240V is the equivalent of putting twenty 12V car batteries together in series.

Secondly, because of the way household electrical energy is generated, it is delivered as an **alternating current** (AC). This means that the electrons in the wire oscillate backwards and forwards in the wire. In comparison, a battery provides **direct current** (DC), which means the electrons travel in one direction only.

Fortunately, most household AC systems can be modelled using simple DC circuits.

Electricity bills

The work done, or energy consumed by a household appliance, is measured in joules. It is the power, P, in watts multiplied by the time, t, in seconds.

$E = Pt$

If you have a look at your home electricity bill, you'll see it is measured in **kilowatt hours** (kW h). This gives a convenient number, without scientific notation, that is easy to write on your bill.

To calculate how much an appliance costs to run, multiply its power consumption in kW by the number of hours it runs for. Then take this number in kW h and multiply it by the cost of electricity per kW h. Worked example 5.6.1 shows how to calculate the cost of running of some household appliances.

Worked example 5.6.1

CALCULATING THE COST OF ELECTRICITY

A 2000 W air conditioner runs for 5 hours. Assume the price for household electricity is 26 cents per kW h. How much would it cost to run this air conditioner for 5 hours?	
Thinking	**Working**
Convert the power consumption of the appliance to kW.	$\frac{2000\text{ W}}{1000} = 2\text{ kW}$
Use the appropriate equation to multiply the power of the appliance in kW by the number of hours it operates.	$E = Pt$ $= 2 \times 5$ $= 10\text{ kW h}$
Multiply the number of kW h by the cost per kW h.	Cost $= 10 \times 0.26$ $= \$2.60$

Worked example: Try yourself 5.6.1

CALCULATING THE COST OF ELECTRICITY

A 2500 W iron is used for 2.5 hours. Assume the price for household electricity is 26 cents per kW h. How much would it cost (to the nearest cent) to use this iron for 2.5 hours? Give your answer correct to 2 decimal places.

ELECTRICAL SAFETY DEVICES

Household electrical wires can carry large amounts of energy. This means that they have the potential to do a lot of harm. The inherent danger associated with the use of electricity can be reduced using various safety devices.

Fuses and circuit breakers

Since wires heat up when current passes through them, there is a limit to how much current the wires in a house or building can safely carry. Household wiring systems are designed to prevent wires from becoming **overloaded**. Appliances that draw a lot of current, such as ovens, hot-water systems and air-conditioners, are put on separate circuits to lights and power points.

Despite these precautions, overloading can still occur, most often due to a **short circuit**. A short circuit occurs when an electric circuit contains very little resistance. This can occur in an electrical appliance when the insulation between the active and neutral wires becomes damaged and these wires are in direct contact. In household circuits, short circuits are always dangerous situations. Large amounts of current means that wires will heat up, causing insulation to melt or catch alight.

PHYSICSFILE

Kilowatt hours and Joules

The energy unit of kW h can be converted to joules (J) by multiplying the number of kW h by 3 600 000. Therefore 1 kW h = 3.6×10^6 J or 3.6 MJ.

This value is simply calculated as follows:

$$1\text{ kW h} = 1000\text{ W} \times 60\text{ minutes} \times 60\text{ seconds}$$
$$= 3\,600\,000\text{ watts} \times \text{seconds}$$
$$= 3.6 \times 10^6\text{ J}$$

An electric current will always take the path of least resistance. The globe in Figure 5.6.5(a) is on because the current has no alternative but to pass through the high resistance of the bulb. The globe in Figure 5.6.5(b) does not work because the closed switch provides a zero-resistance alternative pathway for the current—a short circuit.

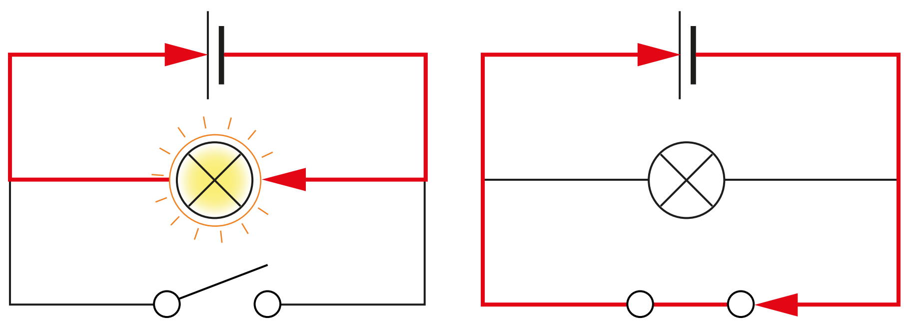


FIGURE 5.6.5 Short circuits. The globe in (a) functions normally, while (b) does not light up.

PHYSICSFILE

Power board overload

Another common reason for circuit overloading is the overuse of power boards and double adaptors. Most power boards are designed to carry a maximum of 7.5 A or 10 A of current. This value is written on them. If too many high-current appliances, such as heaters, kettles and irons, are plugged into a power board, it can overheat. This may cause the insulation around the wires to melt, causing a short circuit or even a fire. Figure 5.6.6 shows the overuse of power boards. This can result in more current being drawn than the maximum rating of the power board that then overloads it resulting in the wires and components overheating and possibly causing a fire.

FIGURE 5.6.6 Overloading of power boards.

Every domestic electric circuit contains either a **fuse** or a **circuit breaker**. The function of both of these components is to interrupt the flow of current if it exceeds a certain value. Unlike a fuse, a circuit breaker can be easily reset after it has been activated, whereas a fuse needs to be physically replaced once it has melted through. Both fuses and circuit breakers can be chosen for different amounts of current, since some appliances such as ovens and hot-water systems might typically draw much larger amounts of current than regular power points.

Earth wires

Many household electrical appliances such as kettles, toasters and ovens have metal cases. If the active wire inside the appliance becomes loose and touches the case, then the whole case becomes electrically live. If anyone touches the case, the current will flow through their body, with possibly fatal consequences.

To prevent this, an earth wire is permanently connected to the metal case of the appliance, as shown in Figure 5.6.7. When the appliance is plugged in, this wire is connected via the household wiring system to the earth. This means that, if the active wire touches the case, a short circuit will be created and current will immediately flow directly to earth. The large amount of current that flows in this situation should trip the fuse or circuit breaker, alerting users of the appliance to the problem.

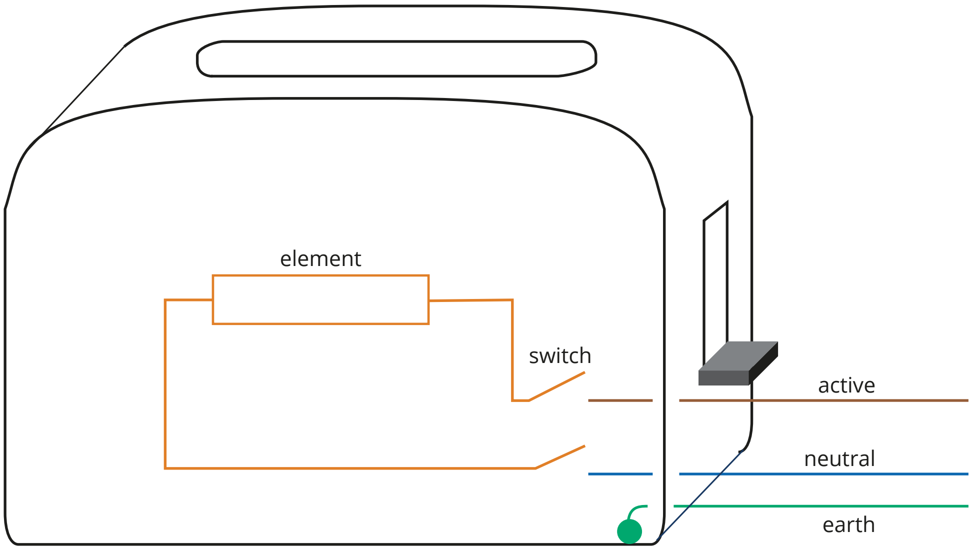


FIGURE 5.6.7 The earth wire inside a metal toaster is permanently connected to the casing.

Double insulation

Double insulation is another way of protecting against the possibility of a loose active wire inside an appliance. This process involves using two insulating barriers to protect users. Often this is done by making the case of the appliance out of plastic. This acts as an insulating layer in addition to the plastic insulation surrounding the active wire inside the appliance. Double-insulated appliances do not need an earth wire, so their electrical plugs only have two pins. They also usually have a special symbol on their cases, indicating that they are double insulated, as shown in Figure 5.6.8.

FIGURE 5.6.8 The double square symbol indicates that this appliance has double insulation.

Residual current devices

Residual current devices, also known as RCDs or earth leakage systems, detect any difference between the current in the active wire and the current in the neutral wire. In a properly operating circuit, these two currents should be exactly the same, but in opposite directions. The most likely reason for a difference between the active and neutral currents is that some current is going to earth through a fault or, in a worse case, through a person. If this happens, the RCD is able to switch off the supply in about 20 milliseconds, hopefully preventing any serious harm.

ELECTRIC SHOCK

Despite all the safety features of modern electrical systems and appliances, each year approximately 50 Australians are killed in electrical accidents. The effect of **electric shock** depends on a number of factors including:

- the amount of current passing through the body
- its duration
- the path it takes through the body.

When attending to a victim of electrocution, it is important to first check if they are still in contact with the electrical source. First aid should only be administered when it is safe to approach the victim (Figure 5.6.9). If it is safe to touch the patient and they are breathing, place them in the recovery position. If they are not breathing then start CPR.

The amount of current that will flow through the body if it is in good contact with a 240 V source is well above the level that can cause death. Anything that improves the contact, like wet hands or bare feet, lowers the resistance and increases the current, potentially to life-threatening levels. Although rubber boots and gloves will increase resistance and lower the current, these should *not* be used as a form of protection in place of other sensible electrical safety precautions.

FIGURE 5.6.9 The recovery position.

Since our bodies are relatively poor conductors, electrical energy passing through our bodies is quickly converted into heat and can cause terrible internal and external burns. Table 5.6.1 shows the likely effect on the human body of a half-second electric shock at different currents.

TABLE 5.6.1 The effect of a half-second electric shock on the body.

Current (mA)	Effect on the body
1	able to be felt
3	easily felt
10	painful
20	muscles paralysed—cannot let go
50	severe shock
90	breathing upset
150	breathing very difficult
200	death likely
500	serious burning, breathing stops, death inevitable

The amount of electrical energy that enters the body depends on the duration of electrocution. This is why the high voltage spark from a Van de Graaff generator is harmless: the duration of the current is about a microsecond and the total energy delivered is tiny. Table 5.6.2 shows the likely effect on the human body of a 50 mA shock for different time periods.

TABLE 5.6.2 The effect of time on the severity of a shock.

Time (s)	Effect on the body
less than 0.2	noticeable but usually not dangerous
0.2–4	significant shock, possibly dangerous
more than 4	severe shock, possible death

The path that the current takes through the body is also important in determining its effect. Since our bodies are controlled by electrical impulses along the nerves, any current that flows into the body from an external source may interfere with our vital functions. In particular, any current flowing from one arm to the other may cause the chest muscles to contract and breathing to stop. Current through the heart regions can cause the muscles to become uncoordinated and the heart function stops. A brief current of about 80 mA is sufficient to cause fibrillation (irregular contraction of the heart muscle) if it flows directly through the heart. This is the cause of most electrical fatalities.

5.6 Review

SUMMARY

- Electrical energy use in the home is usually measured in kilowatt hours, where $1\ \mathrm{kWh} = 1000\ \mathrm{W} \times 1\ \mathrm{h}$
- The effect of electrocution depends on the amount of current passing through the body, its duration and the path it takes through the body.
- The danger associated with the use of electricity in the home can be managed by using fuses, circuit breakers, double insulation and residual current devices.

KEY QUESTIONS

1 How does a fuse or circuit breaker increase household electrical safety?

A by breaking the flow of current if it becomes too high

B by breaking the flow of current if a difference is detected between the currents in the active and neutral wires

C by taking current to the earth if the metal casing of the appliance becomes live

D by providing an extra layer of electrical insulation

2 How does double insulation increase household electrical safety?

A by breaking the flow of current if it becomes too high

B by breaking the flow of current if a difference is detected between the currents in the active and neutral wires

C by taking current to the earth if the metal casing of the appliance becomes live

D by providing an extra layer of electrical insulation

3 Convert 10 kWh into J.

4 One of the values given in the information below is incorrect. Use your knowledge of kWh to determine which value it is.

A 750 W air conditioner uses 0.75 kWh of energy in 1 hour. A typical price for household electricity is 27 cents per kWh. Therefore this air conditioner would cost approximately $10 to run for 5 hours.

5 Why is it that there are only two cables coming into the house from the street and yet power points always have three connections?

6 The function of a fuse is to burn out, and thus turn off the current, if the circuit is overloaded. Why is it always placed in the active wire at the meter box rather than the neutral one, given that this function could be fulfilled if it was in either?

7 What is the function of the 'earth stake' that will normally be found near a meter box?

8 A toaster cable with conductors coloured red, black and green is to be joined to another cable with brown, blue and green/yellow conductors. Peter has joined the red and blue, black and brown, and green and green/yellow. Will the toaster work normally when it is plugged in and turned on? Why is the way he has connected the cables dangerous?

9 An appliance was mistakenly wired between the active and earth instead of between the active and neutral. Explain why that is a very dangerous thing to do, even though the appliance will appear to work normally.

10 How much current would flow through a person with dry hands and a total contact resistance of 100 kΩ when they touch a 240 V live wire?

Chapter review

05

KEY TERMS

alternating current
ammeter
charge
circuit breaker
conductor
conventional current
coulomb
current
direct current
earth
effective resistance
electric current
electric shock
electrical potential energy
electricity
electron flow
elementary charge
fuse
insulators
ion
ionised
junction
kilowatt hour
metals
net charge
non-metals
non-ohmic
ohmic
overload
parallel circuit
potential difference
power
resistance
resistor
residual current device
series circuit
short circuit
transfer
transform
voltmeter
volts

1 Approximately how many electrons make up a charge of –3 C?

2 What will be the approximate charge on 4.2×10^{19} protons?

3 Which charged particles are moving when an electric current flows in a circuit?

A negatively charged electrons
B positively charged electrons
C positively charged protons
D both negative and positive charges

4 An alpha particle consists of two protons and two neutrons. Calculate the charge on an alpha particle.

5 Calculate the current that flows when 0.23 C of charge passes a point in a circuit each minute.

6 Compare the meaning of the terms ‘conventional current’ and ‘electron flow’.

7 A current of 1.6 A flows for 100 seconds. Calculate:

a the amount of charge, in coulombs, that moves past a point in this time
b the number of electrons that move past a point in this time.

8 A current of 0.04 A flows for a certain amount of time. In this time 5×10^{18} electrons move past a point. Calculate:

a the amount of charge, in coulombs, that moves past a point
b the amount of time that the current is flowing.

9 A phone battery has a potential difference of 3.8 V. If 2 C of charge is drawn from the battery, what amount of energy would this provide?

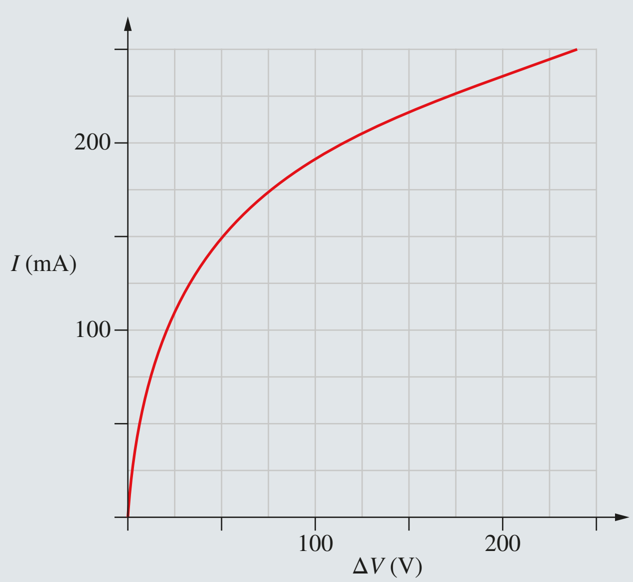
10 A battery does 2 J of work on a charge of 0.5 C to move it from point A to point B. Calculate the potential difference between the two points A and B.

11 How much power does an appliance use if it does 2500 J of work in 30 minutes?

12 A battery gives a single electron 1.4×10^{-18} J of energy. Calculate the potential difference supplied by the battery to two decimal places.

13 A 230 V appliance consumes 2000 W of power. The appliance is left on for 2 hours. What current flows through the appliance?

14 Calculate the resistance at 50 V of the non-ohmic conductor with the I–ΔV characteristics shown in the graph in the figure below.



15 Two resistors, R_1 and R_2, are wired in series. Which of the following gives the equivalent series resistance for these two resistors?

A $R_T = R_1 + R_2$

B $\dfrac{1}{R_T} = \dfrac{1}{R_1} + \dfrac{1}{R_2}$

C $R_T = R_1 - R_2$

D $R_T = R_1 \times R_2$

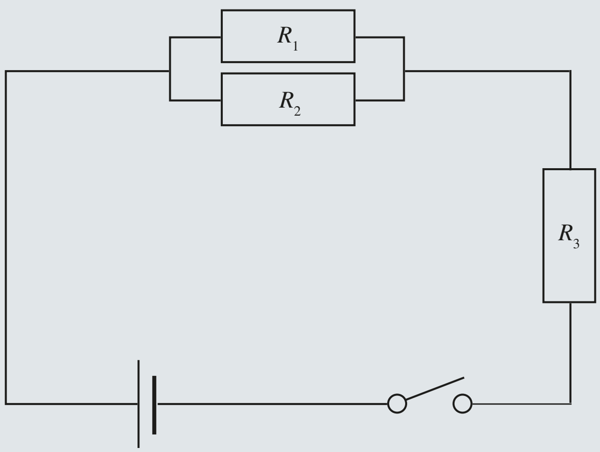
16 An electrical circuit is constructed as shown below. Use the information given about the circuit diagram to answer the following questions.

The electric cell provides 3 V.

The total resistance R_T of the circuit is 8.5 Ω.

R_2 has a resistance of 15 Ω.

The total resistance of resistors R_1 and R_2 is 5.0 Ω.



a Find the value of R_3.

b Find the current through R_3.

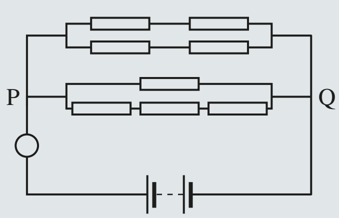
c Find the potential difference across the parallel pair R_1 and R_2.

d Find the current through R_2.

e Find the current through R_1.

f Find the value of R_1.

17 Eight equal-value resistors are connected between points P and Q. The value of each of these resistors is 20.0 Ω.



a The circle in the circuit diagram represents either an ammeter or a voltmeter. Identify which type of meter this should be.

b Calculate the total equivalent resistance of the circuit. Assume that the resistance of the meter and power source can be ignored.

18 Explain how an earth wire improves electrical safety in the home.

19 Sketch a circuit diagram showing how four 10 Ω resistors can be connected using a combination of series and parallel wiring to have a total equivalent resistance of 10 Ω.

20 A circuit consists of a 12 V battery and three 20 Ω light bulbs. The bulbs are initially connected in series.

a Calculate the power output of the circuit.

b The circuit is changed so that the bulbs are connected in parallel. Calculate the power output of the circuit.

c Compare the power drawn in the parallel circuit with that of the series circuit.

21 Which of the following would be most likely to cause serious electrocution harm to a human being?

A a high voltage spark from a Van de Graaff generator; duration = 1 ms

B 3 mA current; duration = 0.5 s

C 50 mA current; duration = 0.1 s

D 50 mA current; duration = 4.5 s

22 A 3 kW heating unit runs for 4 hours. If household electricity costs 30 cents per kW h, how much does it cost to run the heater for this time?

23 The best definition of power from the following choices is:

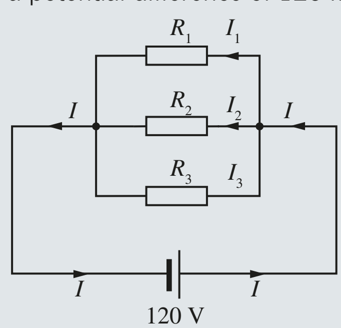
A the total amount of energy consumed by a circuit component

B the current drawn by a circuit component each second

C the rate at which potential difference is supplied to a circuit component

D the rate at which energy is transformed by a circuit component.

24 Consider the following circuit where three resistors R_1, R_2 and R_3 are connected in parallel. Assume that R_1 = 100 Ω, R_2 = 200 Ω and R_3 = 600 Ω. The battery provides a potential difference of 120 V.



a Calculate the total resistance, R_T, in the circuit.

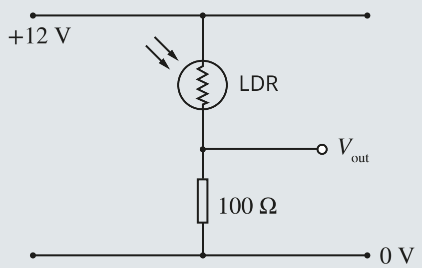
b Calculate the current, I, in the circuit.

c Determine the branch currents I_1, I_2 and I_3.

d What is the power output, P, of the battery?

e Calculate the total power consumed by all of the resistors in the external circuit.

25 In the simple LDR light-detector circuit shown below, ΔV_{out} is to be used to activate an alarm when the ambient light reaches a certain level. At this particular light level, the resistance of the LDR is 200 Ω. The alarm activates whenever ΔV_{out} is above the trigger level.

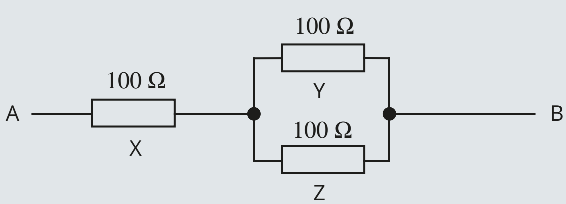


a What is the value of ΔV_{out} at which the alarm should activate?

b Will the alarm activate when the light is above or below the particular level of concern? Explain your answer.

c When it is very dark, what would you expect ΔV_{out} to become?

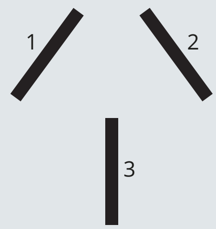
26 Three 100 Ω resistors are connected as shown. The maximum power that can safely be dissipated in any one resistor is 25 W.



a What is the maximum potential difference that can be applied between points A and B?

b What is the maximum power that can be dissipated in this circuit?

27 The diagram below shows the three sockets that can be seen when looking directly at a power point, numbered 1, 2 and 3.



a Which number is the earth socket?

b Which number is the active socket?

c Which number is the neutral socket?

28 Why is the shock received when a finger touches a live wire likely to be less severe than the shock received by a person who touches a live wire with a pair of uninsulated pliers?

29 It is said that a fuse protects property and a safety switch or residual current device protects lives. Explain why this statement is true.

REVIEW QUESTIONS

Section one: Short response

1 A bucket is filled with equal amounts of hot and cold water. The hot water is originally at 80.0°C and the cold water at 10.0°C. Calculate the temperature of the final mixture.

2 State whether the following statements are correct or incorrect. Write an explanation justifying your choice.
  - **a** Water in a bottle covered with wet cloth will be cooler than water in a similar bottle left uncovered in the same environment.
  - **b** Pasta will cook quicker if the water is boiling faster.
  - **c** A patient burned by steam at 100°C sustains significantly more serious injuries than a patient who is scalded by boiling water.

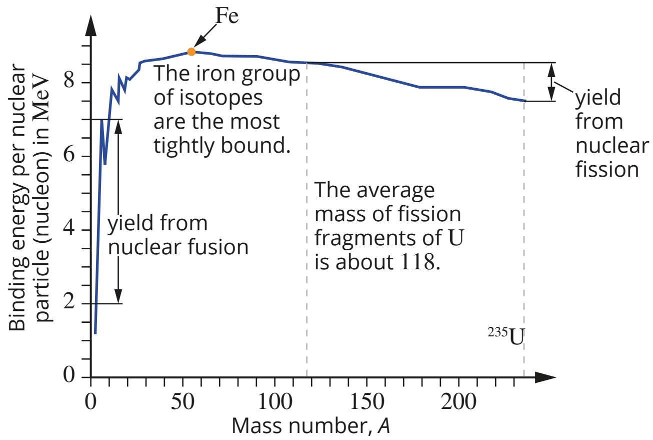
3 Give clear explanations for each of the following.
  - **a** The difference between a fuse and a residual current device in terms of their operation and the kind of protection that they provide.
  - **b** A short circuit in a household electrical appliance may trigger the circuit breaker in its circuit.
  - **c** Some household electrical appliances have plugs with three prongs and some have only two, yet both are safe.

4 Write a balanced nuclear equation for each of the following.
  - **a** A lithium-7 nucleus is bombarded with a high-speed proton resulting in the production of two identical particles.
  - **b** Gold-185 emits an alpha particle.
  - **c** Thallium-218 undergoes beta minus decay.

5
  - **a** Two slow-moving protons are travelling directly towards each other. Will the protons collide and fuse together? In your answer make reference to the forces acting on the protons and the energy barrier.
  - **b** Two fast-moving protons are travelling directly towards each other. The protons collide and fuse together. Discuss the forces that act on the protons and make reference to the energy barrier in your answer.

6 An electron and a positron collide and annihilate each other with the production of two photons.
  - **a** Explain the source of the energy of the photons.
  - **b** Assuming both the electron and positron each had minimal kinetic energy, calculate the combined energy (in joules and MeV) of the photons produced.

7 Use the following graph to answer the questions below.



  - **a** Explain what is meant by binding energy per nucleon.
  - **b** The element iron is considered to be the most stable of all the nuclei. Using the graph, explain why this is the case.
  - **c** Fission of uranium-235 results in fission fragments of average mass number around 118. By referring to the binding energy per nucleon for the fuel and the fragments, explain why there is a net energy release in a fission reaction.

8 A student performs an experiment in which an electric motor is used to lift a 200 g weight through 2.0 m, thus increasing its potential energy by 4.0 J. They will determine the efficiency of the motor from measurements of the rate at which the weight is lifted. In the experiment, when the voltage was 6.0 V a current of 0.25 A was measured and the weight took 5.0 s to rise the 2.0 m. Calculate the efficiency of the motor.

Section two: Problem-solving

9 A chef is considering opening an ice cream parlour that makes ice cream using liquid nitrogen. The liquid nitrogen serves the dual purpose of freezing the water in the ice cream and aerating it at the same time. He intends to make ice cream that is 70% water. He assumes that only the water needs to freeze to set the ice cream.

He collects the following data to use in his calculations:

Boiling temperature of liquid nitrogen at atmospheric pressure: 77.0 K

Specific heat capacity for nitrogen gas: $1.34 \times 10^3\,\text{J kg}^{-1}\,\text{K}^{-1}$

Latent heat of vaporisation of liquid nitrogen: $1.99 \times 10^5\,\text{J kg}^{-1}$

Specific heat capacity of cream and sugar mix: $3.80 \times 10^3\,\text{J kg}^{-1}\,\text{K}^{-1}$

Heat of fusion of water: $3.34 \times 10^5\,\text{J kg}^{-1}$

- **a** How much heat is required to vaporise 1.00 kg of liquid nitrogen at 77.0 K?
- **b** How much heat is absorbed by 1.00 kg of nitrogen gas as it heats from 77.0 K to 273.0 K?
- **c** How much heat must be removed to cool 200.0 g of refrigerated sugar and cream mix at 8.00°C down to make ice cream at 0.00°C?
- **d** What mass of liquid nitrogen does the chef need per 200.0 g ice cream portion if he takes the sugar and cream mix from the fridge at 8.00°C and freezes it at 0.00°C?

10 The slow neutron bombardment of uranium-235 produces many fission products via a series of reactions. One such reaction yields caesium-140 and rubidium-93 as products.

- **a** Write an equation for this fission reaction.
- **b** Calculate the energy, in both electron volts and joules, which would be released from the fission of one uranium-235 atom in this reaction.

Particle	Mass (kg)	Mass (u)
U-235	3.90221×10^{-25}	235.07295
Cs-140	2.32338×10^{-25}	139.96265
Rb-93	1.54301×10^{-25}	92.95241
neutron	1.67493×10^{-27}	1.00899

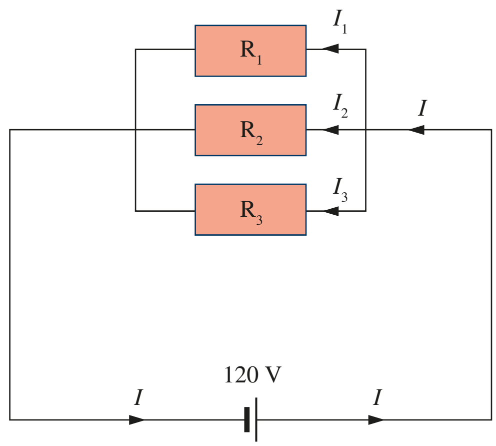
- **c** Calculate the energy in joules that would be released from the fission of 1.00 kg of uranium-235 atoms if the reaction products were only those given above.
- **d** The Trojan nuclear power station in Oregon produces 9.76×10^{13} J of electrical energy per day. In theory, what mass of U-235, in kg, would be needed to produce this energy?
- **e** This power station in fact uses around 2.78 kg of U-235 per day. Discuss the reasons for the difference in the mass of U-235 used.

11 A Geiger counter measures the radioactive disintegrations from a sample of a certain radioisotope. The count rate is recorded in the table.

Activity (Bq)	800	560	400	280	200	140
Time (min)	0	5	10	15	20	25

- **a** Plot a graph of activity versus time.
- **b** Use your graph to estimate the activity of the sample after 13 minutes.
- **c** What is the half-life of this element?
- **d** Determine the activity of the sample after 30 minutes.

12 Three resistors, $R_1 = 100\,\Omega$, $R_2 = 200\,\Omega$ and $R_3 = 600\,\Omega$, in a circuit are connected in parallel with a 120 V battery.



- **a** Calculate the total resistance in the circuit.
- **b** Determine the total current in the circuit.
- **c** Calculate the branch currents I_1, I_2 and I_3.
- **d** Determine the power output of the battery.
- **e** Calculate the total power consumed by all the resistors in the circuit.

Section three: Comprehension

13 Speed skiing

Speed skiing is an extreme winter sport. It is usually conducted at high altitudes where is there is less air so less air resistance. It involves the skier reaching speeds well in excess of 200 km h^{-1} using nothing more than gravity and their ability to minimise friction as they accelerate down a steep slope of 1000 to 1400 m to reach the highest speed they possibly can. They have a slope of length around 400 m to reach maximum speed before passing through a set of light beams 100 m apart that control a timing device. These skiers usually reach their maximum speed well before the first light beam. These light beams enable a very accurate time to be determined for how long the skier took to travel the 100 metres.

The equipment these skiers use is highly specialised, being developed to reduce friction and air resistance as much as possible. Their skis are specially shaped—longer, wider and heavier than those used for normal downhill and recreational skiing—with their extra weight coming not only from their larger size but from the fact they are made of both wood and steel. The extra length and width helps spread the skier's weight force over a larger surface area to help in reducing the frictional forces between the skis and the snow. The shape gives the skis a low profile and, coupled with the weight and its distribution, helps minimise wind resistance and keeps the tip of the skis on the snow.

Another specially designed piece of equipment individually made for each skier is their helmet. This is made to match their body size and the tuck position they adopt as they travel down the slope. It is designed to direct the wind from the top of their head straight down their back.

Speed skiers wear modified boots to allow them to bend their knees and lean well forward as they accelerate down the slope. This lowers their centre of gravity and provides a lower profile, so reducing air resistance. The development of individually customised, polyurethane coated skin-tight suits has also enabled them to minimise air resistance. Both the tightness of fit of these suits and the low interruption offered to airflow by their coating allow air to pass more freely over them. To help smooth the airflow around the calves, a dense foam fitting is made to go inside the suit and extends from the knees to the top of the boots.

To enable them to move off from the starting platform, the skiers use specially designed poles. After they are moving down the slope these poles become important in helping the skier brace their arms into the sides of their bodies. These poles are specially made for each skier and are bent in such a way as to wrap around their body.

For a skier to reach maximum speed in less than 400 m their initial acceleration is critical. The ski slope for speed skiing must be very steep at the start and not start to level out until after the 100 m section where the timing takes place. There is then a lengthy braking region to allow the skier to stop safely.

a What aspect of their motion is being determined by measuring the time the skier takes to travel 100 m?

b Suggest a reason why this system is used to measure speed rather than a radar gun at the start of the 100 m section.

c The statement 'using nothing more than gravity and their ability to overcome friction' is used in relation to the skiers achieving acceleration. Explain how their acceleration is affected by these two factors.

d Why do they reach a maximum speed and not continue to accelerate all the way down the slope?

e Speed skis are wider, longer and heavier than normal skis. How does this help reduce friction?

f List four other design factors stated in the article that bring about reduced friction and state how each one does this.

g The statement is made that 'For a skier to reach maximum speed in less than 400 m their initial acceleration is critical'. Explain why this is the case.

The following information relates to questions h–m.

The current male world speed skating record is 254.958 km h^{-1}, which was achieved by Ivan Origone on 28 March 2016 high in the French Alps. He started at an altitude of 2720 m and came to a halt at an elevation of 2285 m.

h If Origone reached this speed after skiing 340 m down the slope, how many seconds was it from when he left the start?

i Calculate his average acceleration from the start to when he reached his maximum speed.

j Calculate how long it took him to travel between the two light beams in the 100 m section.

k Given his mass is 70.0 kg, calculate his kinetic energy at the start of the 100 m timing section of his run.

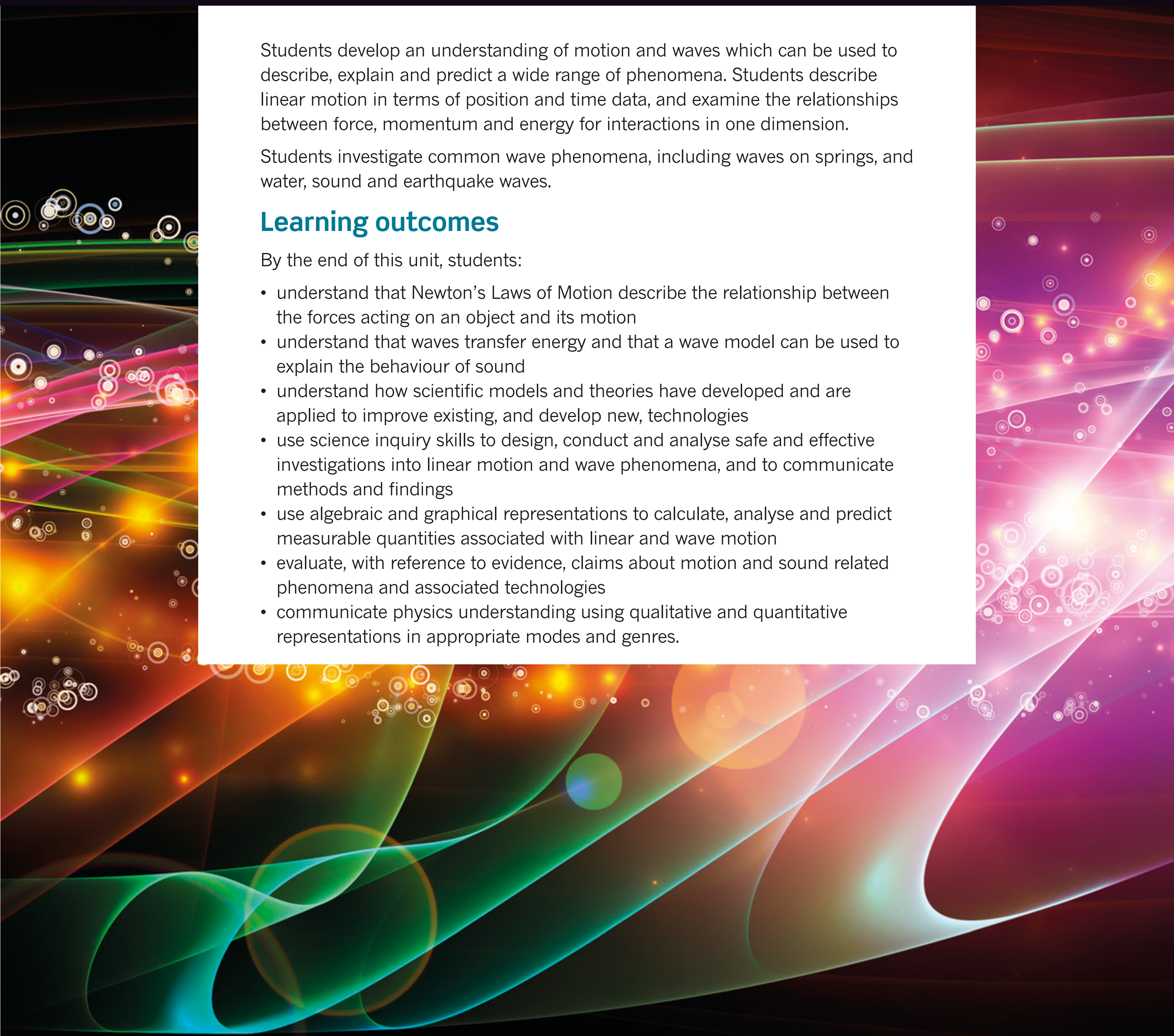
l The energy conversion efficiency from potential to kinetic energy was 93%.

- **i** Calculate his potential energy at the start of his run.
- **ii** At what altitude did he reach his maximum speed?

m Calculate the frictional force acting on him between the start and when he reached his maximum speed.

UNIT 2

Linear motion and waves



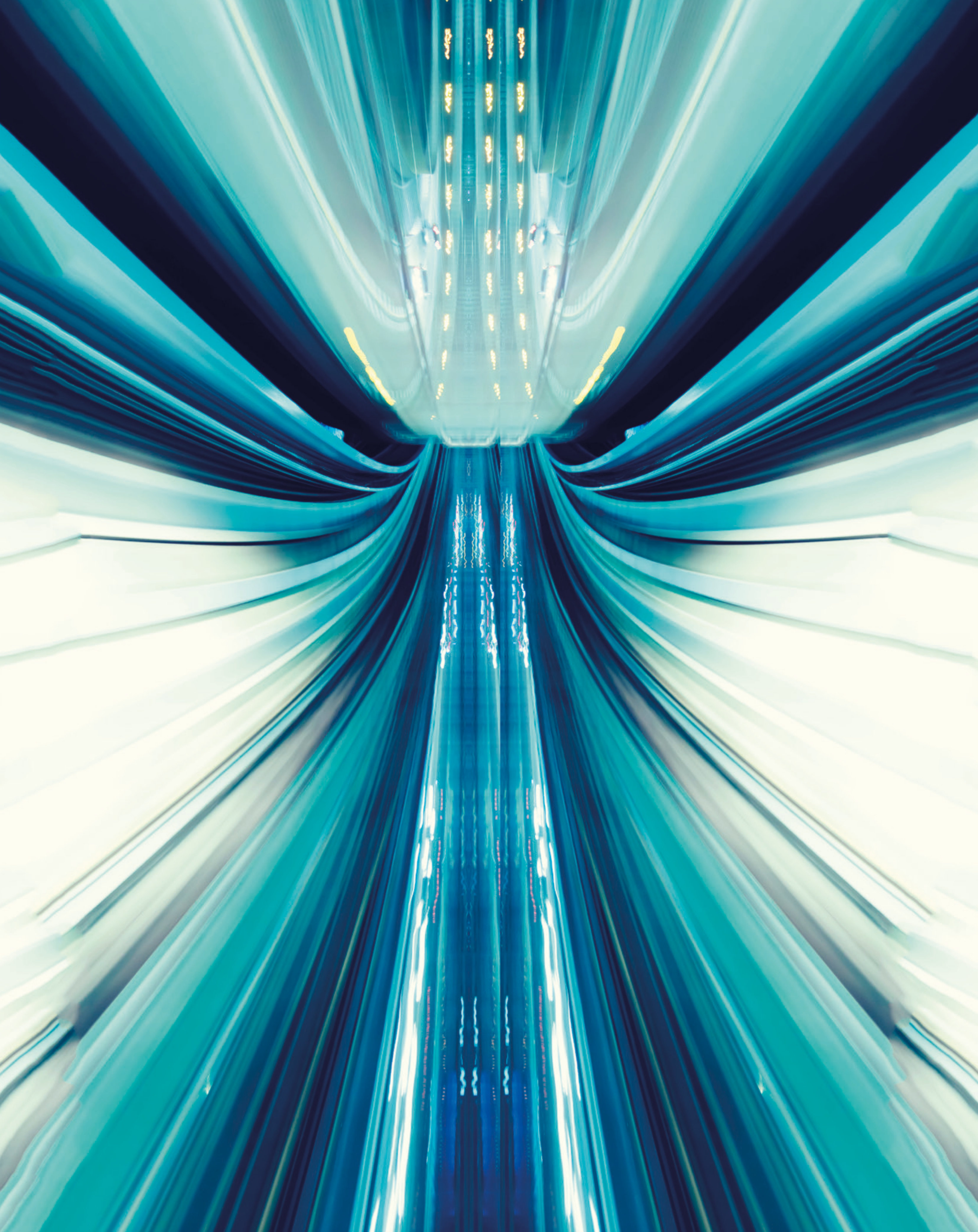
Students develop an understanding of motion and waves which can be used to describe, explain and predict a wide range of phenomena. Students describe linear motion in terms of position and time data, and examine the relationships between force, momentum and energy for interactions in one dimension.

Students investigate common wave phenomena, including waves on springs, and water, sound and earthquake waves.

Learning outcomes

By the end of this unit, students:

- understand that Newton's Laws of Motion describe the relationship between the forces acting on an object and its motion
- understand that waves transfer energy and that a wave model can be used to explain the behaviour of sound
- understand how scientific models and theories have developed and are applied to improve existing, and develop new, technologies
- use science inquiry skills to design, conduct and analyse safe and effective investigations into linear motion and wave phenomena, and to communicate methods and findings
- use algebraic and graphical representations to calculate, analyse and predict measurable quantities associated with linear and wave motion
- evaluate, with reference to evidence, claims about motion and sound related phenomena and associated technologies
- communicate physics understanding using qualitative and quantitative representations in appropriate modes and genres.



CHAPTER 06

Scalars and vectors

Scalars and vectors are mathematical representations of quantities that are used in physics. An understanding of scalars and vectors is essential to learning concepts involving forces and motion.

By the end of this chapter, you will be able to distinguish between scalar and vector quantities. You will be able to use arrows to represent vectors and then add and subtract vectors in one and two dimensions.

Science Understanding

Linear motion and force

- distinguish between vector and scalar quantities, and add and subtract vectors in two dimensions
- representations, including graphs, vectors, and equations of motion, can be used qualitatively and quantitatively to describe and predict linear motion

6.1 Scalars and vectors

You will come into contact with many physical quantities in the natural world every day. For example, time, mass and distance are all physical quantities. Each of these physical quantities has units with which to measure them. For example, seconds, kilograms and metres.

Some measurements only make sense if there is also a direction included. For example, a GPS system tells you when to turn and in which direction. Without both of these two instructions, the information is incomplete.

All physical quantities can be divided into two broad groups based on what information you need for the quantity to make sense. These groups are called **scalars** and **vectors**. Both of these types of measures will be investigated throughout this section.

SCALARS

There are a number of properties in nature that can be measured or determined and described using only a number and unit. For example, if the time taken for a student to travel to school is measured, you need the **magnitude** (size) and the **units** in order to understand the journey. It may take 90 minutes or one and a half hours—the number is important and the units are also important.

Quantities that require magnitude and units are called scalar quantities. Scalars do not need direction.

Examples of scalar quantities are:

- time
- distance
- volume
- speed
- temperature.

VECTORS

Vector quantities require magnitude, units and direction in order to make sense.

Examples of vectors include:

- position
- displacement
- velocity
- acceleration
- force
- momentum.

These measures are discussed in more detail in the coming chapters.

There are also a number of physical quantities that require magnitude, direction and units to make sense. For example, to describe a force applied on an object, you would need to state the magnitude, direction of force, and the unit to fully convey the information.

VECTORS AS ARROWS

A vector is a measurement that has both a magnitude and a direction. A vector can be visually represented as an arrow.

Figure 6.1.1 shows two vector diagrams. In a **vector diagram,** the length of the arrow indicates the magnitude of the vector. The arrowhead shows the direction of the vector. The direction of the vector is always from its tail to its head.

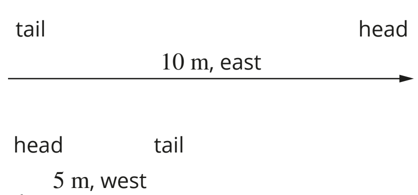


FIGURE 6.1.1 Two vector diagrams. As the top vector is twice as long as the bottom vector, it represents a measure twice the magnitude of the bottom vector. The arrowheads indicate that the vectors are in opposite directions.

A force is a push or a pull and the unit of measure for force is the newton (N). If you push a book to the right, it will respond differently to if you push the book to the left. Therefore, a force is only described properly if a direction is included, and so force is considered to be a vector. Forces are described in more detail in Chapter 8. Force is an important concept to understand in physics, so many of the examples in this chapter refer to forces.

In most vector diagrams, the length of the arrow is drawn to scale so that it accurately represents the magnitude of the vector.

In the scaled vector diagram in Figure 6.1.2, a force $F = 4\,\text{N}$ to the left acting on the toy car is drawn as a 2 cm–long arrow pointing to the left. In this example, a scale of 1 cm = 2 N force is used.

An exact scale for the magnitude is not always used. However, it is important that vectors are drawn relative to one another. For example, a vector of 50 m north should always be about half as long as a vector of 100 m north.

FIGURE 6.1.2 A force of 4 N to the left acts on a toy car.

Point of application of arrows

Vector diagrams may be presented slightly differently depending on what they are depicting. If the vector represents a force, the tail end of the arrow is drawn at the point where the force is applied to the object. If it is a displacement vector, attach the tail of the arrow to the position where the object starts. Friction vectors are drawn at the point where they act between an object and a surface.

Figure 6.1.3 shows a force applied by a foot to a ball (95 N east) and an opposing friction force (20 N west).

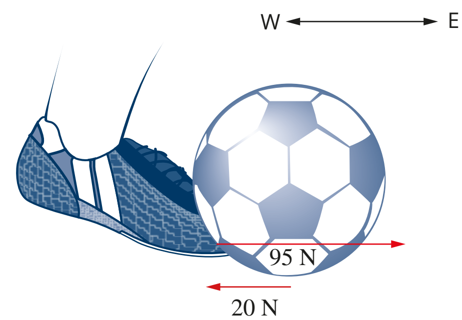


FIGURE 6.1.3 The force on the ball acts at the point of contact between the ball and the foot. The friction force acts between the ball and the ground. The kicking force, as indicated by the length of the arrow, is larger than the friction force.

DIRECTION CONVENTIONS

Vectors need a direction in order to make sense. However, for any description of a vector quantity to be useful, there needs to be a way of describing the direction that everyone understands and agrees upon.

Vectors in one dimension

For vector problems in one **dimension**, there are a number of **direction conventions** that can be used. For example:

- forwards or backwards
- up or down
- left or right.

You can also use more formal conventions such as:

- north or south
- east or west.

As you can see, for vectors in one dimension there are only two directions possible. The two directions must be in the same dimension or along the same line. The direction convention used should be presented graphically in all vector problems. This is shown in Figure 6.1.4. Arrows like these are placed near the vector diagram so that it is clear which convention is being used.

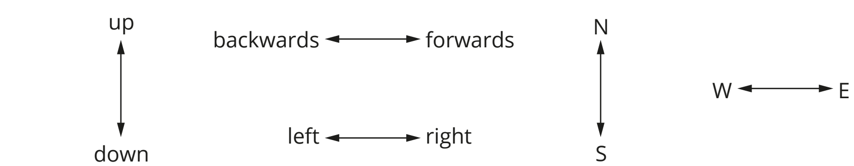


FIGURE 6.1.4 Some common one-dimensional direction conventions.

Sign convention

In calculations involving one-dimensional vectors, a sign convention can also be used to convert physical directions to the mathematical signs of positive and negative. For example, forwards can be positive and backwards can be negative, or right can be positive and left can be negative. A vector of 100 m up can be described as +100 m provided the relationship between sign and direction conventions are clearly indicated in a legend or key. Some examples are provided in Figure 6.1.5.

The advantage of using a sign convention is that the signs of positive and negative can be entered into a calculator, while the words 'up' and 'right' cannot. This is useful when adding vectors together. This will be discussed in the next section.

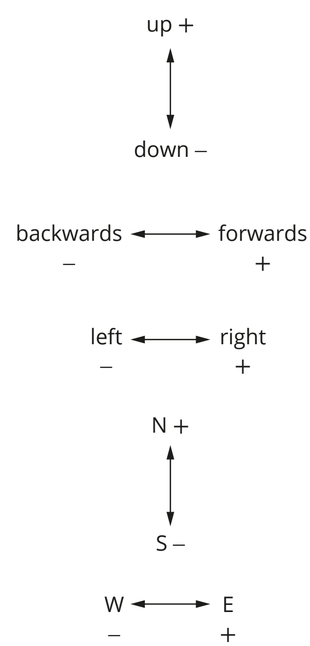
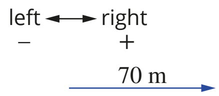


FIGURE 6.1.5 One-dimensional direction conventions can also be expressed as sign conventions.

Worked example 6.1.1

DESCRIBING VECTORS IN ONE DIMENSION



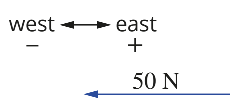
Describe the vector above using:

a the direction convention shown	
Thinking	**Working**
Identify the direction convention being used in the vector.	In this case, the vector is pointing to the right according to the direction convention.
Note the magnitude, unit and direction of the vector.	In this example, the vector is 70 m right.

b an appropriate sign convention.	
Thinking	**Working**
Convert the physical direction to the corresponding mathematical sign.	The physical direction of right is positive and left is negative. In this example, the arrow is pointing right, so the mathematical sign is +.
Represent the vector with a mathematical sign, magnitude and unit.	This vector is +70 m.

Worked example: Try yourself 6.1.1

DESCRIBING VECTORS IN ONE DIMENSION



Describe the vector above using:

a the direction convention shown

b an appropriate sign convention.

Vectors in two dimensions

When vectors are in one dimension, it is relatively simple to understand direction. However, some vectors will require a description in a two-dimensional plane. These planes could be:

- horizontal, which can be defined using north, south, east and west
- vertical, which can be defined in a number of ways including forwards, backwards, up, down, left and right.

The description of the direction of these vectors is more complicated. Therefore, a more detailed convention is needed for identifying the direction of a vector. There are a variety of conventions, but they all describe a direction as an angle from a known reference point.

Horizontal plane

For a horizontal, two-dimensional plane, the two common methods for describing the direction of a vector are:

- full circle (or true) bearing. A 'full circle bearing' describes north as zero degrees true. This is written as 0° T. In this convention, all directions are given as a clockwise angle from north. As an example, 95° T is 95° clockwise from north.
- quadrant bearing. An alternative method is to provide a 'quadrant bearing', where all angles are referenced from either north or south and are between 0° and 90° towards east or west. In this method, 30° T becomes N 30° E, which can be read as 'from north 30° towards the east'.

Using these two conventions, north-west (NW) would be 315° T using a full circle bearing, or N 45° W using a quadrant bearing. Figure 6.1.6 demonstrates these two methods.

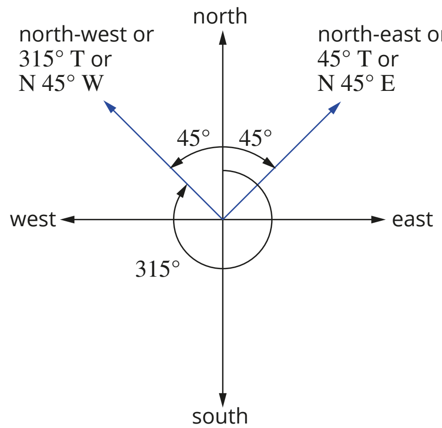


FIGURE 6.1.6 Two horizontal vector directions, viewed from above, using full circle bearings and quadrant bearings.

Vertical plane

For a vertical, two-dimensional plane the directions are referenced to vertical (upwards and downwards) or horizontal (left and right) and are between 0° and 90° up or down. For example, a vector direction can be described as '60° up from horizontal to the left'. The same vector direction could be described as '30° down from the vertical to the left'. The opposite direction to this vector would be '60° down from horizontal to the right'. This example is illustrated in Figure 6.1.7.

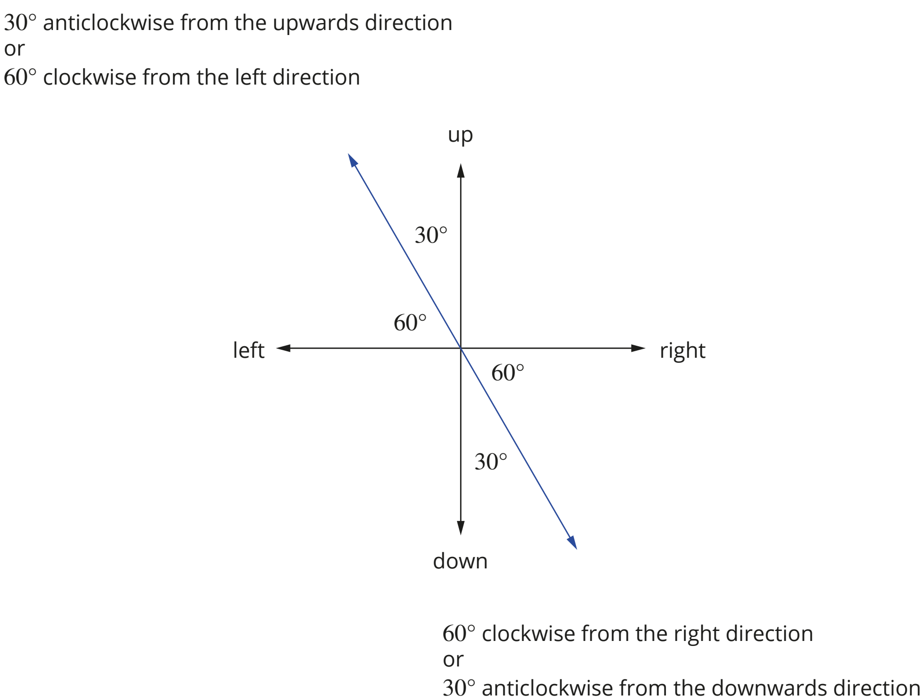


FIGURE 6.1.7 Two vectors in the vertical plane.

PHYSICSFILE

Orienteering

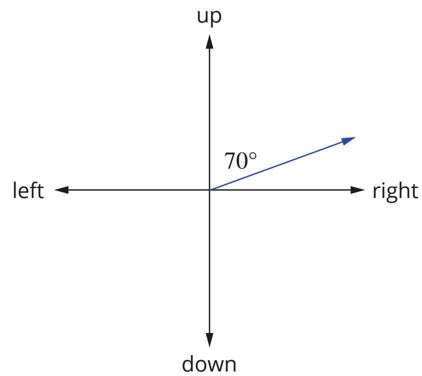
Orienteering is an adventure sport requiring participants to use a compass and map (Figure 6.1.8) to navigate from point to point. Finding your way from place to place involves determining angles measured from any of the cardinal points: north, south, east or west. The vectors between points are shown on the map but contestants need to determine the best way to get there, since the direct line might not be the fastest route. This sport is often timed and competitive and takes place in an unfamiliar, and sometimes challenging, environment. These courses can be followed individually or in teams. Not all orienteering courses have to be completed on foot, as some courses might be designed for mountain bikers or cross-country skiers depending on the environment. There are some permanent orienteering courses in popular spaces in WA such as Whiteman's Park and King's Park, which are free of charge and suitable for any age group. In addition, there are seriously competitive groups that meet on a regular basis.

FIGURE 6.1.8 A compass and map used in orienteering. The numbered points represent the order in which the course should be followed.

Worked example 6.1.2

DESCRIBING TWO-DIMENSIONAL VECTORS

Describe the direction of the following vector using an appropriate method.

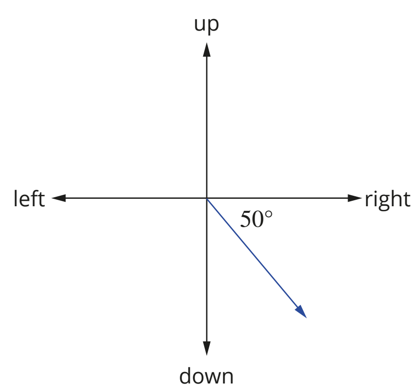


Thinking	Working
Choose the appropriate points to reference the direction of the vector. In this case using the vertical reference makes more sense, as the angle is given from the vertical.	The vector can be referenced to the vertical.
Determine the angle between the reference direction and the vector.	In this example, from vertical to the vector there is 70°.
Determine the direction of the vector from the reference direction.	From vertical, the vector is down to the right.
Describe the vector using the sequence: angle, clockwise or anticlockwise from the reference direction.	This vector is 70° down from vertical to the right.

Worked example: Try yourself 6.1.2

DESCRIBING TWO-DIMENSIONAL VECTORS

Describe the direction of the following vector using an appropriate method.



6.1 Review

SUMMARY

- Scalar quantities require a magnitude and a unit to make sense. No direction is required for scalar quantities.
- Distance, time, speed and mass are examples of scalar quantities.
- Vectors require magnitude, units and direction to make sense.
- Displacement, velocity, acceleration and force are examples of vectors.
- Arrows are used to represent vectors.
- The length of the arrow represents the magnitude of the vector.
- The direction the arrow is pointing indicates the direction of the vector.
- Vector arrows can be drawn to scale, or drawn relative to each other.
- Force vectors are drawn with their tails attached to the point of application on the object.
- Displacement vectors are drawn from the start of the journey towards the end of the journey.
- One-dimensional vectors use direction conventions and sign conventions to describe the direction of the vector. Examples include left and right, up and down, + and –.
- The direction of two-dimensional vectors in the horizontal plane can be described using a full circle bearing or a quadrant bearing. Vectors in the vertical plane can be described using angles measured up or down from the vertical or horizontal, to the right or left.

KEY QUESTIONS

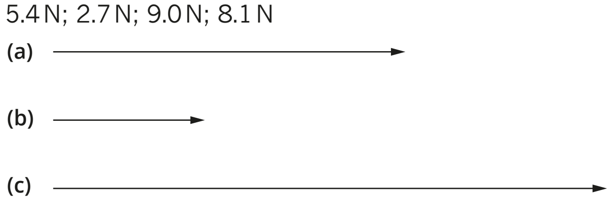
1 What information is required to fully describe a scalar measure?

2 What information is required to fully describe a vector measure?

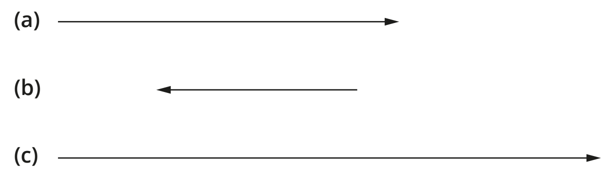
3 Classify each of the following as scalar or vector quantities.

a time
b force
c acceleration
d distance
e position
f displacement
g volume
h momentum
i speed
j velocity
k temperature

4 For the following, decide which of the vector magnitudes provided describes which vector diagram. Note: one of the vector magnitudes is not required.
5.4 N; 2.7 N; 9.0 N; 8.1 N



5 For the following, decide which of the vector magnitudes provided describes which vector diagram. Note: one of the vector magnitudes is not required.
10.8 N; –2.7 N; –5.4 N; 16.2 N

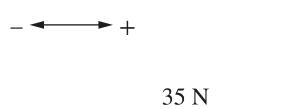


6 Give the opposing direction to each of the following one-dimensional descriptions.

a up
b north
c backwards
d down
e west
f negative

7 Why is it sometimes appropriate to rename direction conventions with a positive or negative sign—for example, + instead of north or – instead of left?

8 Describe the following vector using an appropriate convention.



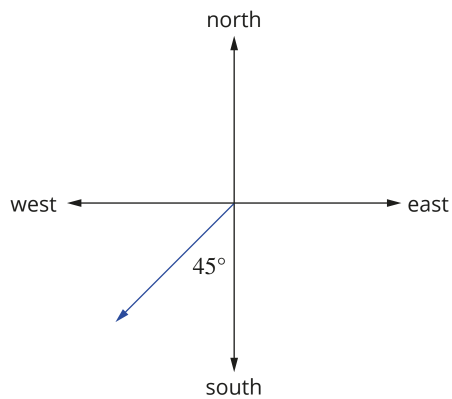
6.1 Review *continued*

9 Describe the following vectors using:

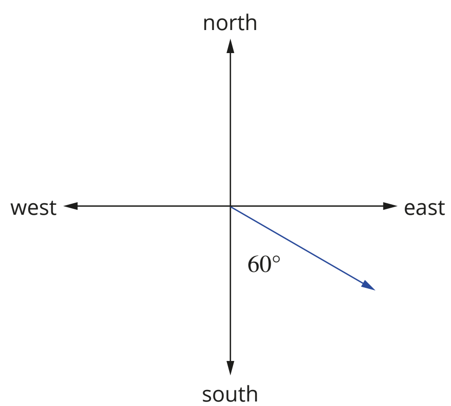
i full circle bearings

ii quadrant bearings.

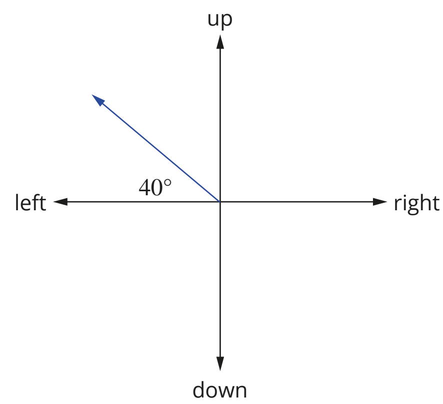
a



b



10 Describe the following vector using appropriate conventions.



6.2 Adding vectors in one and two dimensions

In real situations, more than one vector may act on an object. If this is the case, it is helpful to analyse the associated vector diagrams to find out the overall or combined effect of the vectors.

When vectors are combined, it is called adding vectors. Adding vectors that are in one dimension means finding the resultant vector for a number of vectors that are in the same line. It is also possible to find the resultant of a number of vectors that are in two dimensions. The vectors can be in any direction, as long as they are all in the same plane.

Vectors can also be added in three dimensions, but this is beyond the realm of this course.

ADDING VECTORS IN ONE DIMENSION

When two or more vectors are in the same dimension, it means that the vectors are either pointing in the same direction or in the opposite direction. They are **collinear** (in line with each other). For example, the displacements 10 m west, 15 m east and 25 m west are all in one dimension. They are all in the same or opposite direction to each other.

Graphical method of adding vectors

Vector diagrams, like those shown in Figure 6.2.1, are convenient for adding vectors. To combine vectors in one dimension, draw the first vector, then start the second vector with its tail at the head of the first vector. Continue adding arrows 'head to tail' until the last vector is drawn. The sum of the vectors, or the **resultant** vector, is drawn from the tail of the first vector to the head of the last vector.

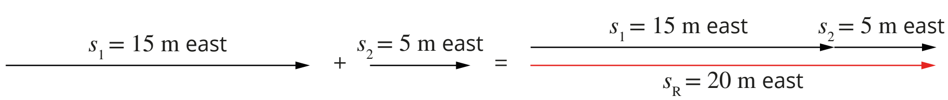


FIGURE 6.2.1 Adding vectors head to tail. This particular diagram represents the addition of 15 m east and 5 m east. The resultant vector, shown in red, is 20 m east.

In Figure 6.2.1, the two vectors s_1 (15 m east) and s_2 (5 m east) are drawn separately. The two vectors are then drawn with s_1 and s_2 connected head to tail. The resultant vector s_R is drawn from the tail of s_1 to the head of s_2. The magnitude (size) of the resultant vector can be deduced from the magnitudes of the separate vectors: 15 m + 5 m = 20 m.

Alternatively, vectors can be drawn to scale, for example: 1 m = 1 cm. The resultant vector is then directly measured from the scale diagram. The direction of the resultant vector is the same as the direction from the tail of the first vector to the head of the last vector.

> Vectors are added head to tail. The resultant vector is drawn from the tail of the first vector to the head of the last vector.

Algebraic method of adding vectors

To add vectors in one dimension algebraically, a sign convention is used to represent the direction of the vectors (see Figure 6.2.2). When applying a sign convention, it is important to provide a key explaining the convention used.

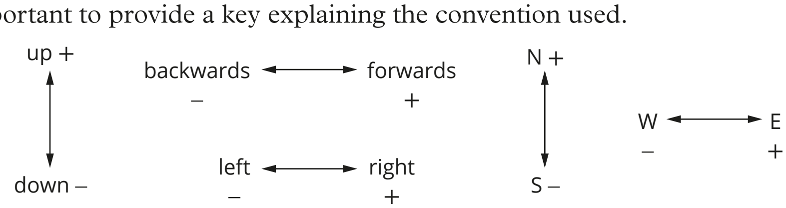


FIGURE 6.2.2 Common sign and direction conventions.

The sign convention allows you to enter the signs and magnitudes of vectors into a calculator. The sign of the final magnitude gives the direction of the resultant vector.

Worked example 6.2.1

ADDING VECTORS IN ONE DIMENSION USING ALGEBRA

Use the sign and direction conventions shown in Figure 6.2.2 on page 181 to determine the resultant vector of a student who walks 25 m west, 16 m east, 44 m west and then 12 m east.

Thinking	Working
Apply the sign conventions to change each of the directions to signs.	25 m west = –25 m 16 m east = +16 m 44 m west = –44 m 12 m east = +12 m
Add the magnitudes and their signs together.	Resultant vector = (–25) + (+16) + (–44) + (+12) = –41 m
Refer to the sign and direction conventions to determine the direction of the resultant vector.	Negative is west. ∴ Resultant vector = 41 m west

Worked example: Try yourself 6.2.1

ADDING VECTORS IN ONE DIMENSION USING ALGEBRA

Use the sign and direction conventions shown in Figure 6.2.2 on page 181 to determine the resultant force on a box that has the following forces acting on it: 16 N up, 22 N down, 4 N up and 17 N down.

ADDING VECTORS IN TWO DIMENSIONS

Adding vectors in two dimensions means that all of the vectors must be in the same plane (coplanar). The vectors can go in any direction within the plane, and can be separated by any angle. The examples in this section illustrate vectors in the horizontal plane, but the same strategies apply to adding vectors in the vertical plane.

The horizontal plane is one that is looked down on from above. Examples include looking at a house plan or map placed on a desk. The direction conventions that suit this plane best are the north, south, east and west convention, or the forwards, backwards, left and right convention. These are shown in Figure 6.2.3.

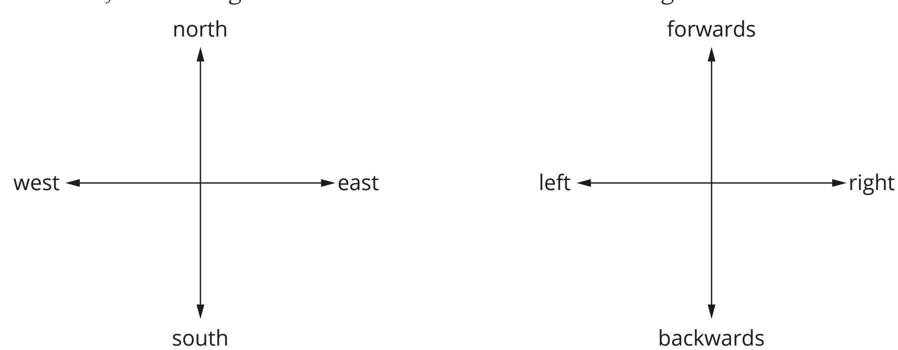


FIGURE 6.2.3 The direction conventions for the horizontal plane.

When two vectors are in the horizontal plane, the angles between them can be right-angled, acute or obtuse.

Graphical method of adding vectors

The magnitude and direction of a resultant vector can be determined by measuring an accurately drawn scaled vector diagram. There are two main ways to do this:

- head to tail method
- parallelogram method.

Head to tail method

To add vectors at right angles to each other using a graphical method, use an appropriate scale and then draw each vector head to tail. The resultant vector is the vector that starts at the tail of the first vector and ends at the head of the last vector. To determine the magnitude and direction of the resultant vector, measure the length of the resultant vector and compare it to the scale, then measure and describe the direction appropriately.

In Figure 6.2.4, two vectors, 30.0 m east and 20.0 m south, are added head to tail. The resultant vector, shown in red, is measured to be about 36 m according to the scale provided. Using a protractor, the resultant vector is measured to be in the direction 34° south of east. This represents a direction of S 56° E when using quadrant bearings.

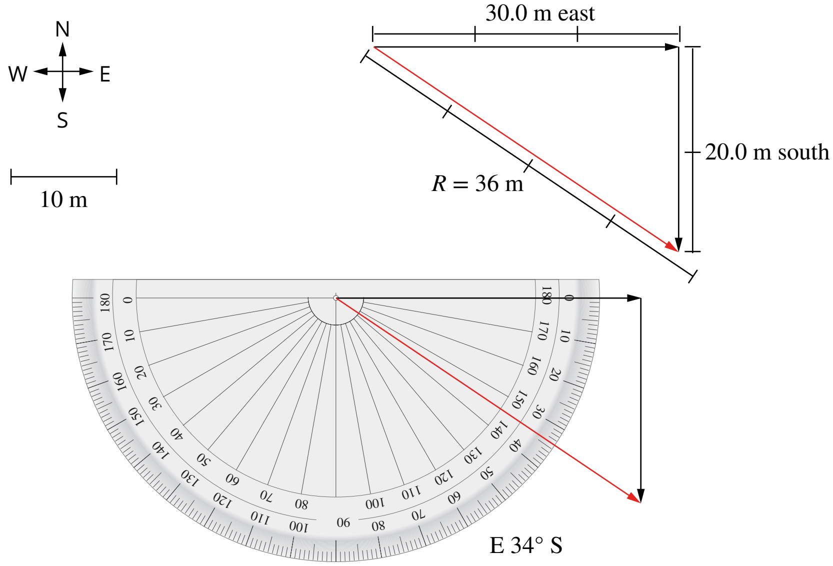


FIGURE 6.2.4 Adding two vectors at right angles, using the graphical method.

If the two vectors are at angles other than 90° to each other, the graphical method is ideal for finding the resultant vector. In Figure 6.2.5, the force vectors 15 N east and 10 N S 45° E are added head to tail. The magnitude of the resultant vector is measured to be about 23 N. The direction of the resultant vector is measured by a protractor from east to be 18° towards the south, which should be written as S 72° E.

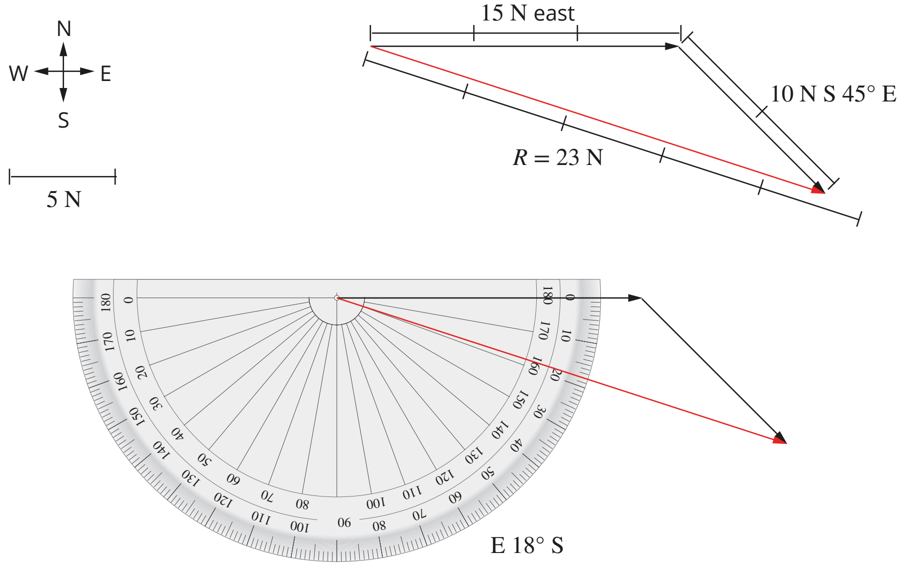


FIGURE 6.2.5 Adding two vectors not at right angles, using the graphical method.

Parallelogram method

An alternative method for determining a resultant vector is to construct a parallelogram of vectors. In this method, the two vectors to be added are drawn tail to tail. Next, a parallel line is drawn for each vector as shown in Figure 6.2.6. In this figure, the parallel lines have been drawn as dotted lines. The resultant vector is drawn from the tails of the two vectors to the intersection of the dotted parallel lines.

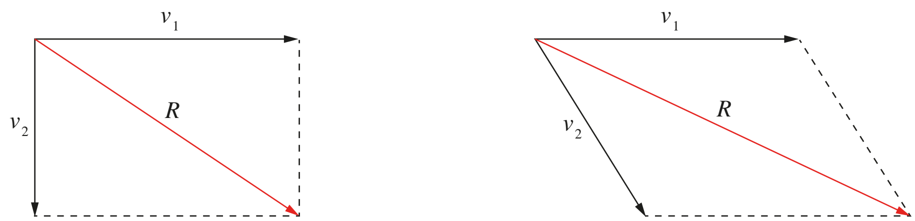
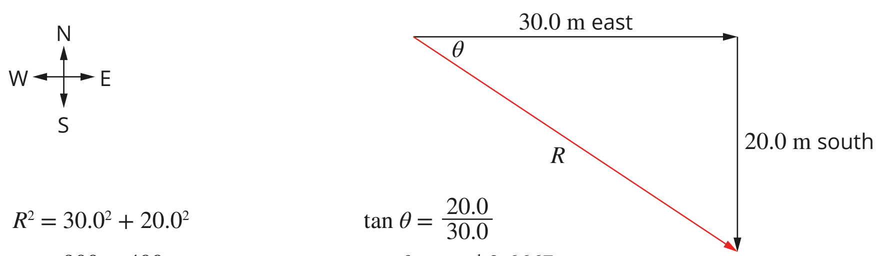


FIGURE 6.2.6 Parallelogram of vectors method for adding two vectors.

Geometric method of adding vectors

Graphical methods of adding vectors in two dimensions only give approximate results as they rely on comparing the magnitude of the resultant vector to a scale and measuring the direction with a protractor. A more accurate method to resolve vectors is to use Pythagoras' theorem and trigonometry. These techniques are referred to as geometric methods. Geometric methods can be used to calculate the magnitude of the vector and its direction. However, Pythagoras' theorem and trigonometry can only be used for finding the resultant vector of two vectors that are at right angles to each other.

In Figure 6.2.7, two vectors, 30.0 m east and 20.0 m south, are added head to tail. The resultant vector, shown in red, is calculated using Pythagoras' theorem to be 36.1 m. The resultant vector is calculated to be in the direction S 56.3° E. This result is more accurate than the answer determined earlier in this section.



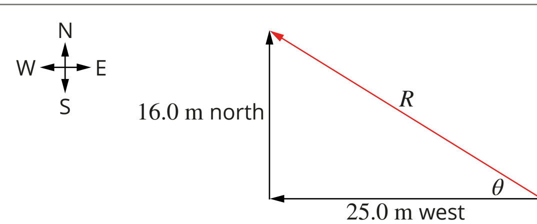
$$R^2 = 30.0^2 + 20.0^2$$
$$= 900 + 400$$
$$R = \sqrt{1300}$$
$$= 36.1 \text{ m}$$

$$\tan\theta = \frac{20.0}{30.0}$$
$$\theta = \tan^{-1} 0.6667$$
$$= 33.7°$$
E 33.7° S (or S 56.3° E)

FIGURE 6.2.7 Adding two vectors at right angles, using the geometric method.

Worked example 6.2.2

ADDING VECTORS IN TWO DIMENSIONS USING GEOMETRY

Determine the resultant vector that represents a child running 25.0 m west and 16.0 m north. Refer to Figure 6.2.2 on page 181 for sign and direction conventions if required.	
Thinking	**Working**
Construct a vector diagram showing the vectors drawn head to tail. Draw the resultant vector from the tail of the first vector to the head of the last vector.	

PHYSICSFILE

Pythagoras' theorem and trigonometric ratios

The Year 11 ATAR Physics syllabus assumes students will be able to:

- use Pythagoras' theorem, similarity of triangles and the angle sum of a triangle
- solve simple sine, cosine and tangent relationships in a right-angle triangle.

Pythagoras' theorem is $a^2 + b^2 = c^2$ where c is the hypotenuse (the longest side) and a and b are the two shorter sides of a right-angled triangle. The hypotenuse is easily recognised as it is directly across from (opposite) the right angle of the triangle.

Most students learn the mnemonic SOHCAHTOA in their maths classes. It is often pronounced soh-cah-toa and provides a way to remember the trigonometric ratios:

$$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}}$$

$$\cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}}$$

$$\tan\theta = \frac{\text{opposite}}{\text{adjacent}}$$

As the two vectors to be added are at 90° to each other, apply Pythagoras' theorem to calculate the magnitude of the resultant vector.	$R^2 = 25.0^2 + 16.0^2$ $= 625 + 256$ $R = \sqrt{881}$ $= 29.7\text{ m}$
Using trigonometry, calculate the angle from the west vector to the resultant vector.	$\tan\theta = \frac{16.0}{25.0}$ $\theta = \tan^{-1} 0.640$ $= 32.6°$
Determine the direction of the vector relative to north or south.	$90° - 32.6° = 57.4°$ The direction is N57.4°W
State the magnitude and direction of the resultant vector.	$R = 29.7\text{ m}$, N57.4°W

Worked example: Try yourself 6.2.2

ADDING VECTORS IN TWO DIMENSIONS USING GEOMETRY

Determine the resultant force when forces of 5.0 N east and 3.0 N north act on a tree. Refer to Figure 6.2.2 for sign and direction conventions if required.

PHYSICS IN ACTION

Surveying

Surveyors use technology to measure, analyse and manage data about the shape of the land and the exact location of landmarks and buildings. They take many measurements, including angles and distances, and use them to calculate more advanced data such as vectors, bearings, co-ordinates, elevations and maps. Surveyors typically use theodolites (see figures 6.2.8 and 6.2.9), GPS survey equipment, laser range finders and satellite images to map the land in three dimensions.

Surveyors are often the first professionals on a building site to ensure that the boundaries of the property are correct. They also ensure that the building is built in the correct location. Surveyors must liaise closely with architects both before and during a building project as they provide position and height data for walls and floors.

FIGURE 6.2.8 Surveying the land with a theodolite.

FIGURE 6.2.9 Surveying equipment being used on a building site.

6.2 Review

SUMMARY

- Combining vectors is known as adding vectors.
- One-dimensional vector addition refers to vectors in a line, while two-dimensional vector addition refers to vectors on a plane.
- Adding vectors in one dimension can be done graphically by using vector diagrams. After adding vectors head to tail, the resultant vector can be drawn from the tail of the first vector to the head of the last vector.
- Adding vectors in one dimension can be done algebraically by applying a sign convention. Vectors with direction become vectors with either positive or negative signs.
- Adding vectors in two dimensions can be estimated graphically with a scale and a protractor.
- An alternate method of adding vectors in two dimensions is to construct a parallelogram of vectors.
- Adding vectors in two dimensions can be calculated using Pythagoras' theorem and the trigonometric ratios of a right-angled triangle.

KEY QUESTIONS

1 Every day, Monday to Friday, a man drives his car 3.0 km east to work and returns the same route home. He drives nowhere else.

a What distance will he have driven in a week?

b What is his displacement at the end of a week?

2 Add the following vectors to find the resultant vector: 3 m up, 2 m down and 3 m down.

3 Determine the resultant vector of a toy train that is made to move in these directions: 23 m forwards, 16 m backwards, 7 m forwards and 3 m backwards.

4 When adding vector B to vector A using the head to tail method, from what point, and to what point, is the resultant vector drawn?

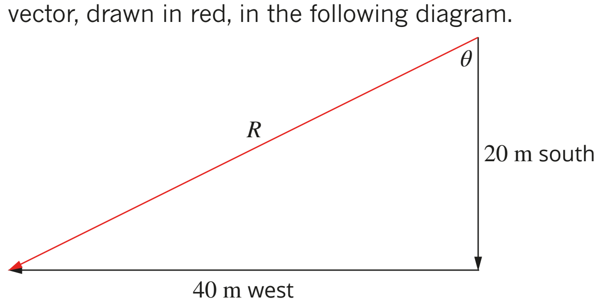
A from the head of A, to the tail of B

B from the tail of B, to the head of A

C from the head of B, to the tail of A

D from the tail of A, to the head of B

5 Describe the magnitude and direction of the resultant vector, drawn in red, in the following diagram.



6 Forces of 2000 N north and 6000 N east act on an object. What is the resultant force acting?

7 Calculate the magnitude of the resultant vector when 30.0 m south and 40.0 m west are added?

8 Determine the resultant force acting on an object being pulled north by force 3000 N; south by 5000 N and east by 5000 N.

9 Determine the resultant vector of the following combination to: 3350 N forward, 6220 N backwards, 2235 N forwards and 634 N forwards.

6.3 Subtracting vectors in one and two dimensions

The previous section discussed combining or adding vectors. In physics, there are times when the difference between two vectors has to be determined. For example, a change in velocity is determined by the final velocity minus the initial velocity. In other words, you must subtract vectors. One way to subtract one vector from another is to add the opposite vector.

SUBTRACTING VECTORS IN ONE DIMENSION

To find the difference between two vectors, you must subtract the initial vector from the final vector. To do this, work out which is the initial vector, then reverse its direction. These two vectors are then added: the final vector and the opposite of the initial vector.

This technique can be applied both graphically and algebraically.

Graphical method of subtracting vectors

Velocity is a quantity that gives an indication of how fast an object is moving in a certain direction. It is a vector because the direction is important when stating the velocity of an object. For example, the velocity of the tennis ball moving towards the racquet in Figure 6.3.2 is different from the velocity of the tennis ball as it leaves the racquet. The concept of velocity is covered in more detail in Chapter 7, but it is useful to use the example of velocity now when discussing the subtraction of vectors. The processes applied to the subtraction of velocity vectors works for all other vectors.

FIGURE 6.3.2 As velocity is a vector, direction is important. The tennis ball has a different velocity when it leaves the racquet from when it travelled towards the racquet.

To subtract velocity vectors in one dimension using a graphical method, determine which vector is the initial velocity and which is the final velocity. The final velocity is drawn first. The initial velocity is then drawn, but in the opposite direction to its original form. The sum of these vectors, or the resultant vector, is drawn from the tail of the final velocity to the head of the reversed initial velocity. This resultant vector is the difference between the two velocities, or Δv.

In Figure 6.3.3, the two separate velocity vectors v_1 ($9\,\text{m}\,\text{s}^{-1}$ east) and v_2 ($3\,\text{m}\,\text{s}^{-1}$, east) are drawn separately. The initial velocity, v_1, is then drawn again in the opposite direction: $-v_1$ or $9\,\text{m}\,\text{s}^{-1}$ west.

FIGURE 6.3.3 Subtracting vectors using the graphical method: the initial vectors.

PHYSICSFILE

Double negatives

When a negative number is multiplied by another negative number, the result is a positive number. It is also illustrated when a negative number is subtracted from another number. The effect is to add the two numbers together. For example, (5) – (–2) = 7.

It is important to differentiate between the terms subtract, minus, take away or difference between and the term negative. The terms subtract, minus, take away or difference between are processes, like add, multiply and divide. You will find these processes grouped together on your calculator. The term negative is a property of a number that means that it is opposite to positive. There is a separate button on your calculator (usually shown as ±) for this property.

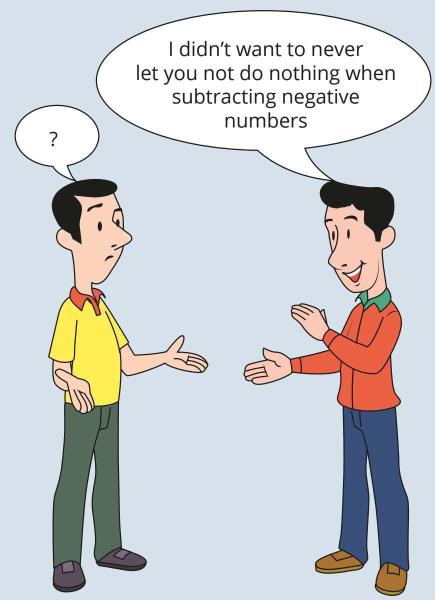


FIGURE 6.3.1 Double negatives can be confusing.

To find the difference between or change in vectors, subtract the initial vector from the final vector. Vectors are subtracted by adding the negative of a vector.

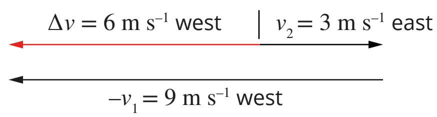


FIGURE 6.3.4 Subtracting vectors using the graphical method: the resultant vector.

Figure 6.3.4 illustrates how the difference between the vectors is found. Firstly, the final velocity, v_2, is drawn. Then the opposite of the initial velocity, $-v_1$, is drawn head to tail. The resultant vector, Δv, is drawn from the tail of v_2 to the head of $-v_1$.

The magnitude of the resultant vector, Δv, can be calculated from the magnitudes of the two vectors. Alternatively, you could draw the vectors to scale and then measure the resultant vector against that scale—for example 1 m s^{-1} = 1 cm.

The direction of the resultant vector, Δv, is the same as the direction from the tail of the final velocity, v_2, to the head of the opposite of the initial velocity, $-v_1$.

Algebraic method of subtracting vectors

To subtract velocity vectors in one dimension algebraically, a sign convention is used to represent the direction of the velocities. Some examples of one-dimensional directions include east and west, north and south and up and down. These options are replaced by positive (+) or negative (–) signs when calculations are performed. To change the direction of the initial velocity, simply change the sign from positive to negative or from negative to positive.

The equation for finding the change in velocity is:

change in velocity = final velocity – initial velocity

$\Delta v = v_2 - v_1$

$\Delta v = v_2 + (-v_1)$

change in velocity = final velocity + the opposite of the initial velocity

The final velocity is added to the opposite of the initial velocity. Since the change in velocity is a vector, it will consist of a sign and a magnitude. The sign of the answer can be compared with the sign and direction convention (Figure 6.3.5) to determine the direction of the change in velocity.

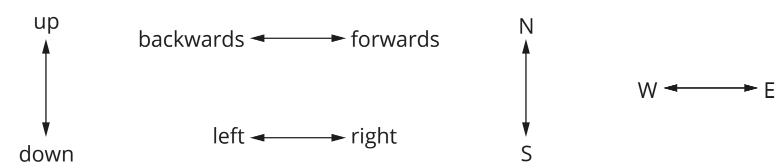


FIGURE 6.3.5 Sign and direction conventions.

Worked example 6.3.1

SUBTRACTING VECTORS IN ONE DIMENSION USING ALGEBRA

Use the sign and direction conventions shown in Figure 6.3.5 to determine the change in velocity of a plane as it changes from 255 m s^{-1} west to 160 m s^{-1} east.	
Thinking	**Working**
Apply the sign and direction convention to change the directions to signs.	v_1 = 255 m s^{-1} west = –255 m s^{-1} v_2 = 160 m s^{-1} east = +160 m s^{-1}
Reverse the direction of the initial velocity, v_1, by reversing the sign.	$-v_1$ = 255 m s^{-1} east = +255 m s^{-1}
Use the formula for change in velocity to calculate the magnitude and the sign of Δv.	$\Delta v = v_2 + (-v_1)$ = (+160) + (+255) = +415 m s^{-1}
Refer to the sign and direction convention to determine the direction of the change in velocity.	Positive is east $\therefore \Delta v$ = 415 m s^{-1} east

Worked example: Try yourself 6.3.1

SUBTRACTING VECTORS IN ONE DIMENSION USING ALGEBRA

Use the sign and direction conventions shown in Figure 6.3.5 to determine the change in velocity of a rocket as it changes from $212\,m\,s^{-1}$ up to $2200\,m\,s^{-1}$ up.

SUBTRACTING VECTORS IN TWO DIMENSIONS

Changing velocity in two dimensions can occur when turning a corner. For example, walking at $3\,m\,s^{-1}$ west, then turning to travel at $3\,m\,s^{-1}$ north. Although the magnitude of the velocity is the same, the direction is different.

A change in velocity in two dimensions can be determined using either the graphical method or the geometric method described in the previous section. The initial velocity must always be reversed before it is added to the final velocity.

The two-dimensional direction conventions were introduced in the previous section and are shown here in Figure 6.3.6.

When a change in a vector occurs, the magnitude and/or the direction of the vector can change.

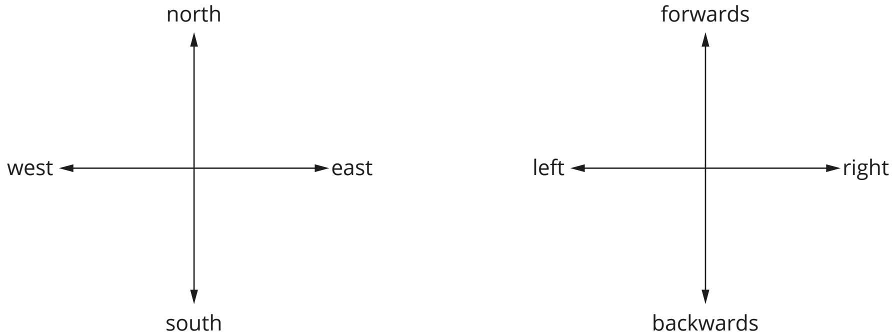


FIGURE 6.3.6 The direction conventions for the horizontal plane.

Graphical method of subtracting vectors

To subtract vectors using a graphical method, use a direction convention and a scale and draw each vector.

Using velocity as an example, the steps to do this are as follows:

- Draw in the final velocity first.
- Draw the opposite of the initial velocity head to tail with the final velocity vector.
- Draw the resultant change in velocity vector, starting at the tail of the final velocity vector and ending at the head of the opposite of the initial velocity vector.
- Measure the length of the resultant vector and compare it to the scale to determine the magnitude of the change in velocity.
- Measure an appropriate angle to determine the direction of the resultant vector.

Figure 6.3.7 shows the velocity vectors for travelling $3\,m\,s^{-1}$ west and then turning and travelling $3\,m\,s^{-1}$ north. The opposite of the initial velocity is drawn as $3\,m\,s^{-1}$ east.

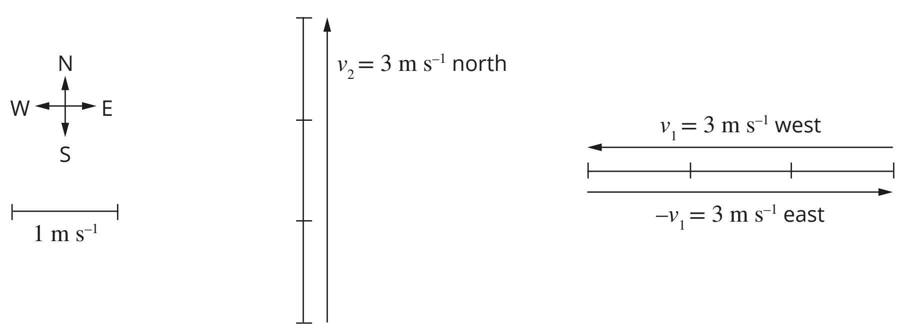


FIGURE 6.3.7 Subtracting two vectors at right angles, using the graphical method.

To determine the change in velocity, the final velocity vector is drawn first, then from its head the opposite of the initial velocity is drawn. This is shown in Figure 6.3.8. The magnitude of the change in velocity (the resultant vector) is shown in red. It is measured to be about 4.3 m s^{-1} according to the scale provided. Using a protractor, the resultant vector is measured to be in the direction N 45° E.

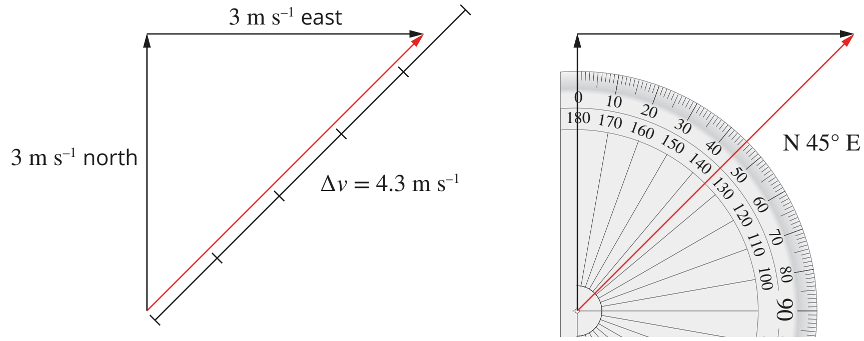


FIGURE 6.3.8 Subtracting two vectors at right angles, using the graphical method.

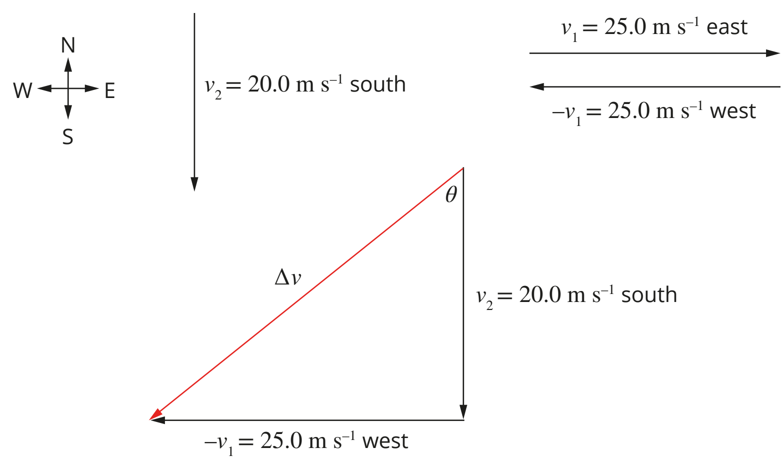
Geometric method of subtracting vectors

The graphical method of subtracting vectors in two dimensions only gives approximate results, as it relies on comparing the magnitude of the change in velocity vector to a scale and measuring its direction with a protractor.

As you saw earlier in the chapter for adding vectors, a more accurate method to subtract vectors is to use Pythagoras' theorem and trigonometry.

Figure 6.3.9 shows how to calculate the resultant velocity when changing from 25 m s^{-1} east to 20.0 m s^{-1} south. The initial velocity of 25.0 m s^{-1} east and the final velocity of 20.0 m s^{-1} south are drawn. Then the opposite of the initial velocity is drawn as 25.0 m s^{-1} west. The final velocity vector is drawn first, then from its head the opposite of the initial velocity is drawn. The resultant velocity vector, shown in red, is calculated to be 32.0 m s^{-1}. The resultant vector is calculated to be in the direction S 51.3° W.

The resultant vector is 32.0 m s^{-1} S 51.3° W.



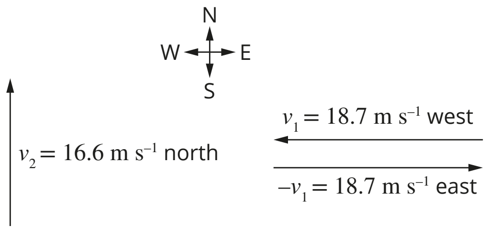
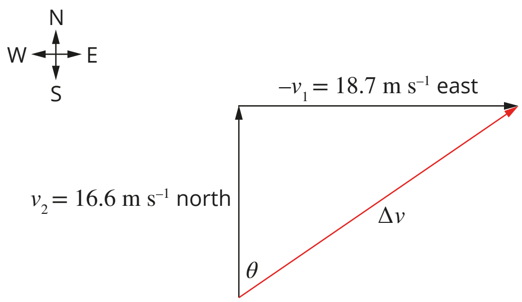
$$R^2 = 25.0^2 + 20.0^2$$
$$= 625 + 400$$
$$R = \sqrt{1025}$$
$$= 32.0 \text{ m s}^{-1}$$

$$\tan\theta = \frac{25.0}{20.0}$$
$$\theta = \tan^{-1} 1.25$$
$$= 51.3°$$
S 51.3° W

FIGURE 6.3.9 Subtracting two vectors at right angles, using the geometric method.

Worked example 6.3.2

SUBTRACTING VECTORS IN TWO DIMENSIONS USING GEOMETRY

Determine the change in velocity of Clare's scooter as she turns a corner if she approaches it at 18.7 m s^{-1} west and exits at 16.6 m s^{-1} north.	
Thinking	**Working**
Draw the final velocity vector, v_2, and the initial velocity vector, v_1, separately. Then draw the initial velocity in the opposite direction.	
Construct a vector diagram drawing v_2 first and then from its head draw the opposite of v_1. The change of velocity vector is drawn from the tail of the final velocity to the head of the opposite of the initial velocity.	
As the two vectors to be added are at 90° to each other, apply Pythagoras' theorem to calculate the magnitude of the change in velocity.	$R^2 = 16.6^2 + 18.7^2$ $= 275.26 + 349.69$ $R = \sqrt{625.25}$ $= 25.0\ \text{m s}^{-1}$
Calculate the angle from the north vector to the change in velocity vector.	$\tan\theta = \dfrac{18.7}{16.6}$ $\theta = \tan^{-1} 1.16$ $= 48.4°$
State the magnitude and direction of the change in velocity.	$\Delta v = 25.0\ \text{m s}^{-1}$ N 48.4° E

Worked example: Try yourself 6.3.2

SUBTRACTING VECTORS IN TWO DIMENSIONS USING GEOMETRY

Determine the change in velocity of a ball as it bounces off a wall. The ball approaches at 7.0 m s^{-1} south and rebounds at 6.0 m s^{-1} east.

6.3 Review

SUMMARY

- To find the difference between, or change in vectors, subtract the initial vector from the final vector.
- Vectors are subtracted by adding the negative, or opposite, of a vector.
- Vector subtraction in one or two dimensions can be determined graphically using a scale and a protractor.
- Vector subtraction in one dimension can be determined algebraically.
- Vector subtraction in two dimensions can be determined geometrically using Pythagoras' theorem and trigonometry.

KEY QUESTIONS

1 A car that was initially travelling at a velocity of $3\,m\,s^{-1}$ west is later travelling at $5\,m\,s^{-1}$ east. What is the difference between the two vectors?

2 Determine the change in velocity of a runner who changes from running on grass at $4\,m\,s^{-1}$ to the right to running in sand at $2\,m\,s^{-1}$ to the right.

3 A student throws a ball up into the air at $4\,m\,s^{-1}$. A short time later the ball is travelling back downwards to hit the ground at $3\,m\,s^{-1}$. Determine the change in velocity of the ball during this time.

4 Tom hits a tennis ball against a wall. If the ball travels towards the wall at $35.0\,m\,s^{-1}$ north and rebounds at $32.5\,m\,s^{-1}$ south, calculate the change in velocity of the ball.

5 Jamelia applies the brakes on her car and changes her velocity from $22.2\,m\,s^{-1}$ forwards to $8.20\,m\,s^{-1}$ forwards. Calculate the change in velocity of Jamelia's car.

6 A jet plane makes a turn after taking off, changing its velocity from $345\,m\,s^{-1}$ south to $406\,m\,s^{-1}$ west. Calculate the change in the velocity of the jet.

7 Yvette hits a golf ball that strikes a tree and changes its velocity from $42.0\,m\,s^{-1}$ east to $42.0\,m\,s^{-1}$ north. Calculate the change in the velocity of the golf ball.

8 A yacht tacks during a race, changing its velocity from $7.05\,m\,s^{-1}$ south to $5.25\,m\,s^{-1}$ west. Calculate the change in the velocity of the yacht.

9 A cyclist travelling north at $40\,km\,h^{-1}$ does a U-turn at the halfway point of a race and slows to $25\,km\,h^{-1}$. Determine:

a the change in speed

b the change in velocity.

10 A motorcyclist travelling south turns around a roundabout and exits at the third exit (now heading west) maintaining a speed of $30.0\,km\,h^{-1}$. Determine his change in velocity.

6.4 Vector components

Sections 6.2 and 6.3 explored how vectors can be combined to find a resultant vector. In physics, there are times when it is useful to break one vector up into two vectors that are at right angles to each other. For example, if a force vector is acting at an angle up from horizontal, as shown in Figure 6.4.1, this vector can be considered to consist of two independent vertical and horizontal components.

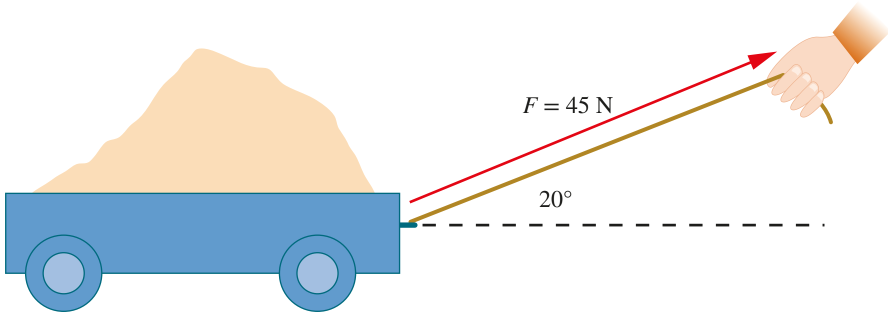


FIGURE 6.4.1 The pulling force acting on the cart has a component in the horizontal direction and a component in the vertical direction.

The components of a vector can be found using trigonometry.

FINDING PERPENDICULAR COMPONENTS OF A VECTOR

Vectors at an angle are more easily dealt with if they are broken up into perpendicular **components**, that is, two components that are at right angles to each other. These components, when added together, give the original vector. To find the components of a vector, a right-angled triangle is constructed with the original vector as the hypotenuse. This is shown in Figure 6.4.2. The hypotenuse is always the longest side of a right-angled triangle and is opposite the 90° angle. The other two sides of the triangle are each shorter than the hypotenuse and form the 90° angle with each other. These two sides are the perpendicular components of the original vector.

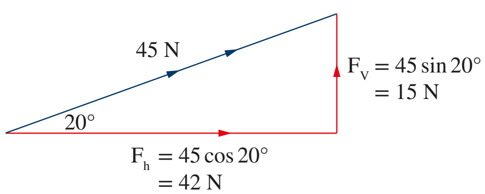


FIGURE 6.4.2 The perpendicular components (shown in red) of the original vector (shown in blue). The original vector is the hypotenuse of the triangle.

Geometric method of finding vector components

The geometric method of finding the perpendicular components of vectors is to construct a right-angled triangle using the original vector as the hypotenuse. This was illustrated in Figure 6.4.2. The magnitude and direction of the components are then determined using trigonometry. A good rule to remember is that no component of a vector can be larger than the vector itself. In a right-angled triangle, no side is longer than the hypotenuse. The original vector must be the hypotenuse and its components must be the other two sides of the triangle.

Figure 6.4.3 shows a force vector of 50.0 N (drawn in black) acting on a box in a direction 30.0° up from horizontal to the right. The horizontal and vertical components of this force must be found in order to complete further calculations.

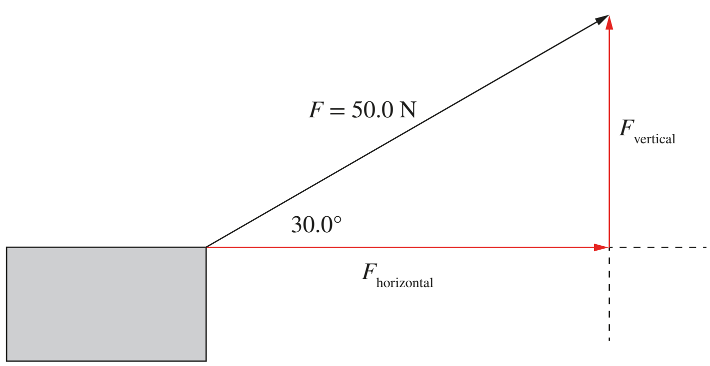


FIGURE 6.4.3 Finding the horizontal and vertical components of a force vector.

PHYSICSFILE

Projectile motion

When objects undergo projectile motion, such as when a ball leaves a tennis racquet, the vertical and horizontal motion are affected differently. So, problem solving begins with separating the initial velocity of the projectile into horizontal and vertical components. At all times during the flight of the ball, the force of gravity is pulling down on the ball, causing it to accelerate downwards until it hits something: another racquet, the net, or the ground. However, the ball will continue to move in the horizontal direction at the same constant velocity as there is no force in that direction, once the ball has left the racquet (figures 6.4.4 and 6.4.5).

FIGURE 6.4.4 A tennis ball in motion accelerates downwards while maintaining a constant horizontal velocity.

FIGURE 6.4.5 By judging the perfect vertical and horizontal components of the velocity required, a tennis player can hit the perfect shot.

The horizontal component vector is drawn from the tail of the 50.0 N vector towards the right, with its head directly below the head of the original 50.0 N vector. The vertical component vector is drawn from the head of the horizontal component to the head of the original 50.0 N vector.

Using trigonometry, the horizontal component of the force is calculated to be 43.3 N horizontally to the right. The vertical component is calculated to be 25.0 N vertically upwards. The calculations are shown below:

$$\cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}}$$

$$\text{adj} = \text{hyp}\cos\theta$$

$$F_h = (50.0)(\cos 30.0°)$$

$$= 43.3\,\text{N horizontal to the right}$$

$$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}}$$

$$\text{opp} = \text{hyp}\sin\theta$$

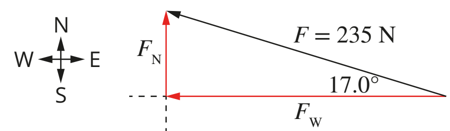
$$F_V = (50.0)(\sin 30.0°)$$

$$= 25.0\,\text{N vertically upwards}$$

Worked example 6.4.1

CALCULATING THE PERPENDICULAR COMPONENTS OF A FORCE

Use the direction conventions to determine the perpendicular components of a 235 N force acting on a bike at a direction of 17.0° north of west.

Thinking	Working
Draw F_W from the tail of the 235 N force along the horizontal direction, then draw F_N from the horizontal vector to the head of the 235 N force.	
Calculate the west component of the force F_W using $\cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}}$	$\cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}}$ adj = hyp cos θ $F_W = (235)(\cos 17.0°)$ = 224.7 N west
Calculate the north component of the force F_N using $\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}}$	$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}}$ opp = hyp sin θ $F_N = (235)(\sin 17.0°)$ = 68.7 N north

Worked example: Try yourself 6.4.1

CALCULATING THE PERPENDICULAR COMPONENTS OF A FORCE

Use the direction conventions to determine the perpendicular components of a 3540 N force acting on a trolley at a direction of 26.5° down from horizontal to the left.

6.4 Review

SUMMARY

- A vector can be resolved into two perpendicular component vectors.
- Perpendicular component vectors are at right angles to each other.
- Any component vectors must be smaller in magnitude than the original vector.
- The hypotenuse of a right-angled triangle is the longest side of the triangle and the other two sides are each smaller than the hypotenuse.
- A right-angled triangle vector diagram can be drawn with the original vector as the hypotenuse and the perpendicular components drawn from the tail of the original to the head of the original.
- The perpendicular components can be determined using trigonometry.

KEY QUESTIONS

1 Rayko applies a force of 462 N on the handle of a mower in a direction of 35.0° down from horizontal to the right.
   a What is the downwards force applied?
   b What is the horizontal right force applied?

2 A force of 25.9 N acts in the direction of S 40.0° E. Find the perpendicular components of the force.

3 A ferry is transporting students to Rottnest Island. At one point in the journey the ferry travels at 18.3 m s^{-1} N 75.6° W. Calculate its velocity in the northerly direction and in the westerly direction at that time.

4 Zehn walks 47.0 m in the direction of S 66.3° E across a hockey field. Calculate the change in Zehn's position down the field and across the field during that time.

5 A cargo ship has two tugs attached to it by ropes. One of the tugs is pulling directly north, while the other tug is pulling directly west. The pulling forces of the tugboats combine to produce a total force of 235 000 N in a direction of N 62.5° W. Calculate the force that each tug boat applies to the cargo ship.

6 Resolve the following forces into their perpendicular components around the north–south line. In part d, use the horizontal and vertical directions.
   a 100 N S 60° E
   b 60 N north
   c 300 N 160°T
   d 3×10^5 N 30° upwards from horizontal to the right.

7 What are the horizontal and vertical components of a 300 N force that is applied along a rope used to drag an object across a yard at 60° up from horizontal to the left?

8 A ball is hit from a racquet with a velocity of 30.0 m s^{-1} at 50.0° up from horizontal to the right. Calculate the horizontal and vertical components of the velocity.

9 A person walks 340 m in a SE direction. How far south has she travelled in this time?

10 A sprinter starts a race by pushing against the starting block with a force of 400 N. If the block is positioned at 70.0° up from horizontal to the left, what horizontal force does the sprinter apply to the block?

Chapter review

06

KEY TERMS

collinear
component
dimension
direction conventions
magnitude
resultant
scalar
units
vector
vector diagram

1 Select the scalar quantities in the list below. (There may be more than one correct answer.)
A force
B time
C acceleration
D mass

2 Select the vector quantities in the list below. (There may be more than one correct answer.)
A displacement
B distance
C volume
D velocity

3 A basketballer applies a force with their hand to bounce the ball. Describe how a vector can be drawn to represent this situation.

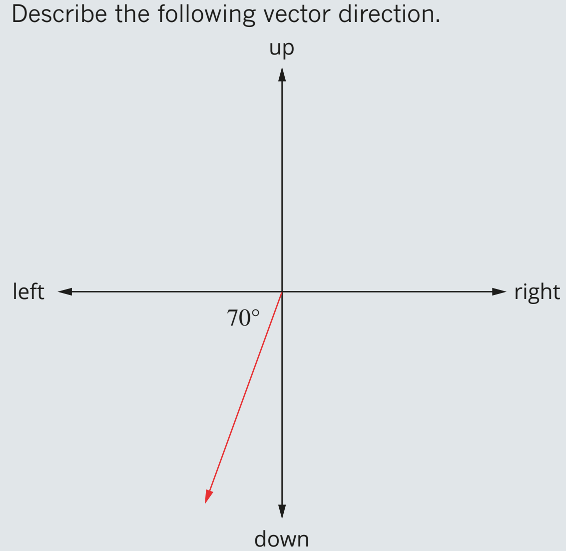
4 Vector arrow A is drawn twice the length of vector arrow B. What does this mean?

5 A car travels 15 m s^{-1} north and another travels 20 m s^{-1} south. Why is a sign convention often used to describe vectors like these?

6 When finding the change in velocity between an initial velocity of 34.0 m s^{-1} south and a final velocity of 12.5 m s^{-1} east, which two vectors need to be added together?

7 If the vector 20 N forwards is written as –20 N, how would you write a vector representing 80 N backwards?

8 Describe the following vector direction.



9 Add the following force vectors using a number line: 3 N left, 2 N right, 6 N right. Then also draw and describe the resultant force vector.

10 Determine the resultant vector of the following combination: 45.0 m forwards, 70.5 m backwards, 34.5 m forwards, 30.0 m backwards.

11 Find the vector which results from the addition of 36 m south and 55 m west.

12 Add the following vectors: 481 N north and 655 N east. Give answers to three significant figures.

13 Determine the change in velocity of a bird that changes from flying 3 m s^{-1} to the right to flying 3 m s^{-1} to the left.

14 A car makes a turn, changing its velocity from 13.0 m s^{-1} south to 18.7 m s^{-1} west. Calculate the change in the velocity vector, Δv, of the car, to three significant figures.

15 Bill hits a cricket ball so that it changes its velocity from 38.8 m s^{-1} east to 55.5 m s^{-1} north. Calculate the change in the velocity vector, to three significant figures.

16 A force of 45.5 N acts in the direction of S 60.0° E. Find the eastern and southern components of this force. Give your answers to three significant figures.

17 Calculate the vertical velocity of a cannonball which is shot at 50.0° up from the horizontal ground to the right at a speed of 400 m s^{-1}.

18 Findlay pulls a heavy load with a force of 200 N at 60.0° up from horizontal to the right and Dougie pulls twice as hard with a different rope at 50.0° up from horizontal to the right. Determine the total horizontal force, which is pulling the load along.

19 Josh shoots a basket 5.0 m away. It just manages to go down into the basket with a vertical velocity of 3.00 m s^{-1} at 20.0° down from vertical to the right. Calculate its actual speed when it goes through the loop.

20 Sandy does a ski jump off a ramp and lands with a speed of 10.0 m s^{-1} at 45.0° down from horizontal to the right. Calculate his vertical and horizontal velocities when he lands.

CHAPTER

07 Linear motion

Motion, from the simple to the complex, is a fundamental part of everyday life. The motion of a gymnast performing a floor routine is a complex form of motion. An Olympic snowboarder competing in a half-pipe event also exhibits a complex form of motion. Simpler examples include a skier travelling in a straight line down a ski run, a train pulling into a station and a swimmer completing a lap of a pool.

Science Understanding

Linear motion and force

- distinguish between vector and scalar quantities, and add and subtract vectors in two dimensions
- uniformly accelerated motion is described in terms of relationships between measurable scalar and vector quantities, including displacement, speed, velocity and acceleration
  - This includes applying the relationships

$$v_{av} = \frac{s}{t},\ a = \frac{v-u}{t},\ v = u + at,\ s = ut + \frac{1}{2}at^2,\ v^2 = u^2 + 2as$$

- representations, including graphs, vectors, and equations of motion, can be used qualitatively and quantitatively to describe and predict linear motion
- vertical motion is analysed by assuming the acceleration due to gravity is constant near Earth's surface.

7.1 Displacement, speed and velocity

In order to analyse and communicate ideas about motion, it is important to understand the terms used to describe motion—even in its simplest form. In this section you will learn about some of the terms used to describe straight-line motion, such as position, distance, displacement, speed and velocity.

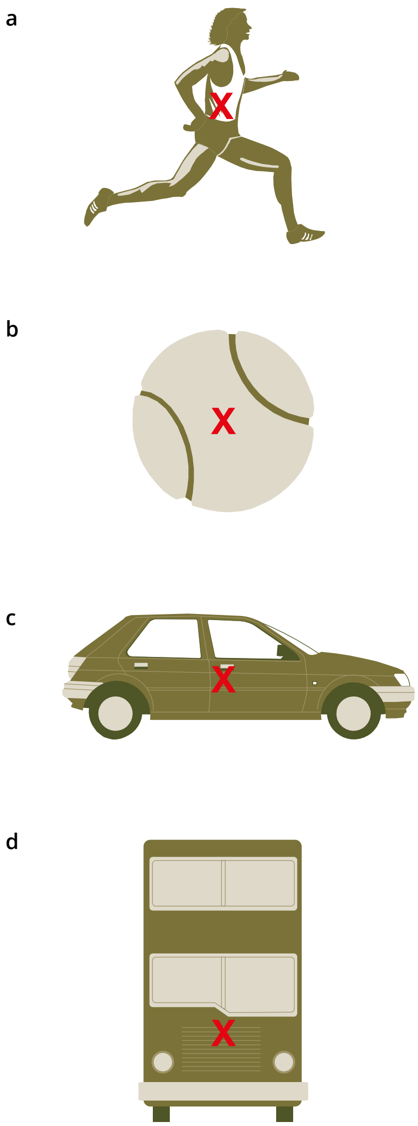


FIGURE 7.1.1 The centre of mass of each object is indicated by a cross.

CENTRE OF MASS

When analysing motion, things are often more complicated than they first appear. As a freestyle swimmer travels at a constant speed of $2\,\text{m}\,\text{s}^{-1}$, their head and torso move forwards at this speed. The motion of their arms, however, is more complex. At times their arms move forwards through the air faster than $2\,\text{m}\,\text{s}^{-1}$, and at other times they are actually moving backwards through the water.

It is beyond the scope of this course to analyse such a complex motion. However, the motion of the swimmer can be simplified by treating the swimmer as a simple object located at a single point called the **centre of mass** or centre of gravity. The centre of mass is the balance point of an object. For a person, the centre of mass is located near the waist. The centres of mass of some everyday objects are shown in Figure 7.1.1. The concept of centre of mass and centre of gravity is discussed in more detail in Year 12.

POSITION, DISTANCE AND DISPLACEMENT

Position

One important term to understand when analysing straight line motion is **position**.

- Position describes the location of an object at a certain point in time with respect to the origin.
- Position is a vector quantity and therefore requires a direction.

Consider a swimmer, Sophie, doing laps in a 50 m pool, as shown in Figure 7.1.2. To simplify her motion, Sophie is treated as a simple point object. The pool can be treated as a one-dimensional number line, with the starting block as the origin. The direction to the right of the starting block is taken to be positive.

Sophie's position as she is warming up behind the starting block in Figure 7.1.2a is −10 m. The negative sign indicates the direction from the origin, i.e. to the left. Her position could also be given as 10 m to left of the starting block.

At the starting block (Figure 7.1.2b), Sophie's position is 0 m, then after swimming half a length of the pool she is +25 m or 25 m to the right of the origin (Figure 7.1.2c).

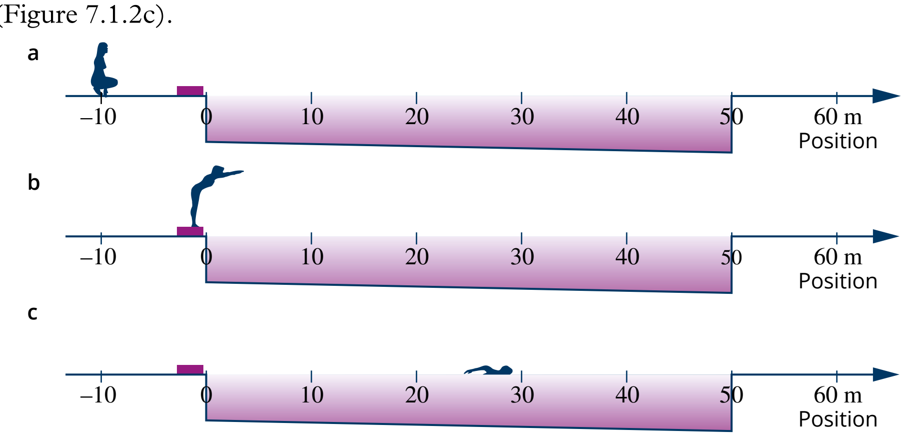


FIGURE 7.1.2 The position of the swimmer is given with reference to the starting block. (a) While warming up, Sophie is at −10 m. (b) When she is on the starting block, her position is zero. (c) After swimming for a short time, she is at a position of +25 m.

Distance travelled

Position describes where an object is at a certain point in time. However, **distance travelled** is how far a body travels during a journey.

- Distance travelled, *d*, describes the length of the path covered during an object's entire journey.
- Distance travelled is a scalar quantity and is measured in metres (m).

For example, if Sophie completes three lengths of the pool, the distance travelled during her swim will be 50 + 50 + 50 = 150 m.

The distance travelled is not affected by the direction of the motion. That is, the distance travelled by an object always increases as it moves, regardless of its direction. The tripmeter or odometer of a car or bike measures distance travelled.

Displacement

Displacement, *s*, is defined as the *change in position* of an object. Displacement takes into account only where the motion starts and finishes. The route taken between these points has no effect on displacement. The sign of the displacement indicates the direction in which the position has changed from the start to the end.

- Displacement is the change in position of an object in a given direction.
- Displacement s = final position – initial position.
- Displacement is a vector quantity and is measured in metres (m).

Consider the example of Sophie completing one length of the pool. During her swim, the distance travelled is 50 m. Her final position is +50 m and her initial position is 0 m. Her displacement is:

s = final position – initial position
= 50 – 0
= +50 m or 50 m in a positive direction

Notice that **magnitude**, units and direction are required for a vector quantity. The distance will be equal to the magnitude of displacement only if the body is moving in a straight line and does not change direction. If Sophie swims two lengths, her distance travelled will be 100 m: 50 m out and 50 m back. However, her displacement during this swim will be:

s = final position – initial position
= 0 – 0
= 0 m

Even though Sophie has swum 100 m, her displacement is zero because the initial and final positions are the same.

The above formula for displacement is useful if you already know the initial and final positions of a body's motion. An alternative method to determine total displacement, if you know the displacement of each section of the motion, is to add up the individual displacements for each section of motion.

total displacement = sum of individual displacements

It is important to remember that displacement is a vector and so, when adding displacements, you must obey the rules of vector addition (discussed in Chapter 6).

In the example above, in which Sophie completed two laps, overall displacement could have been calculated by adding the displacement of each lap:

s = sum of displacements for each lap
= 50 m in the positive direction + 50 m in the negative direction
= 50 + (–50)
= 0 m

PHYSICS IN ACTION

Timing and false starts in athletics

Until 1964, all timing of events at the Olympic Games was recorded by handheld stopwatches (Figure 7.1.3). The reaction times of the judges meant an uncertainty of 0.2 s for any measurement. An electronic quartz timing system introduced in 1964 improved accuracy to 0.01 s, but in close finishes the judges still had to wait for a photograph of the finish before they could announce the places.

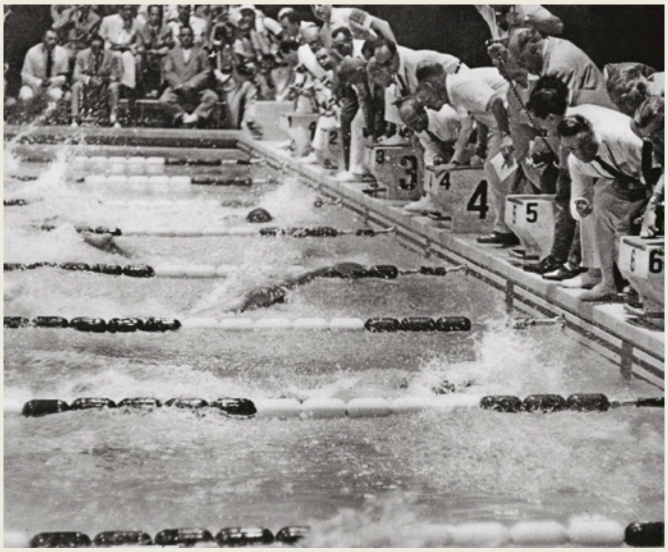


FIGURE 7.1.3 Using stopwatches to time a race in Rome in 1960.

The current timing system used in athletics is a vertical line-scanning video system (VLSV). Introduced in 1991, this electronic timing system is completely automatic. The starting pistol triggers a computer to begin timing. At the finish line, a high-speed video camera records the image of each athlete and indicates the time at which each one crosses the line. This system enables the times of all athletes in the race to be precisely measured to one-thousandth of a second.

Another feature of this system is that it indicates when a runner 'breaks' at the start of the race. Each starting block is connected by electronic cables to the timing computer and a pressure sensor indicates if a runner has left the blocks early (Figure 7.1.4). A reaction time of 0.10 s has been incorporated into the system since 2002. This ensures that a runner has not anticipated the pistol. It also means that a runner can still commit a false start even if their start was after the pistol. A start that is less than 0.10 s after the pistol is registered as false.

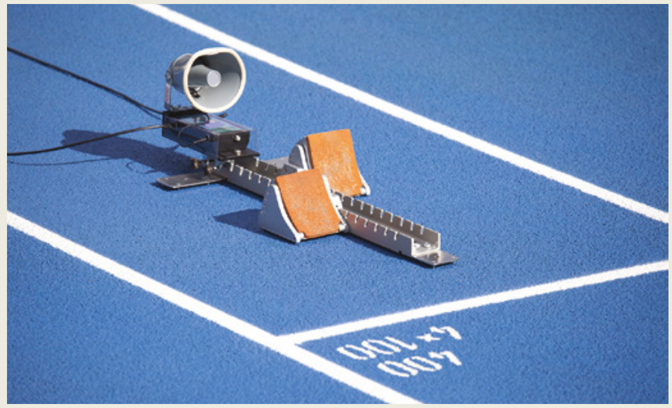


FIGURE 7.1.4 Starting blocks are fitted with pressure sensors to detect false starts.

The most controversial false start of recent times occurred at the athletics World Championships in 2003. It was a quarter-final heat of the men's 100 metres. US runner Jon Drummond was in lane 4 and Asafa Powell of Jamaica was in lane 5. Australia's Patrick Johnson was in lane 6. There had already been a false start in this heat and, since 2002, the rule for false starts in athletics events had been that after one false start, the next athlete to break is disqualified. The athletes went under starter's orders a second time and again there was a false start. The officials examined the computer read-out from the pressure pads on the blocks and determined that both Drummond and Powell were to be disqualified. Asafa Powell immediately left the track. Jon Drummond protested his innocence and proceeded to sit and lie down on the track for the next 20 minutes (Figure 7.1.5). He was widely criticised for his actions, but an analysis of the pressure pad readings revealed that he may have been a little unlucky.

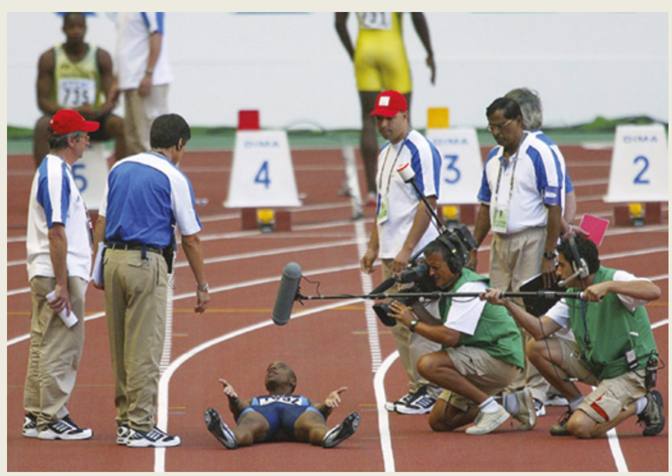


FIGURE 7.1.5 Athlete Jon Drummond protested his false start by lying down on the track.

SPEED AND VELOCITY

For thousands of years, humans have tried to travel at ever greater speeds. This desire has contributed to the development of all sorts of competitive activities, as well as major advances in engineering and design. World records for some of these pursuits are given in Table 7.1.1.

TABLE 7.1.1 World record speeds for a variety of sports or modes of transport (as of June 2017).

Activity	World record speed ($m\,s^{-1}$)	World record speed ($km\,h^{-1}$)
luge	45.6	164
train	168	603
tennis serve	73.1	263
waterskiing (barefoot)	68.3	246
cricket delivery	44.7	161
racehorse	19.7	71

- Speed is defined in terms of the rate at which the distance is travelled. Like distance, speed is a scalar. A direction is not required when describing the speed of an object.
- Velocity is defined in terms of the rate at which displacement changes, and so is a vector quantity. A direction should always be given with a velocity.
- The SI unit for speed and velocity is metres per second ($m\,s^{-1}$), but kilometres per hour ($km\,h^{-1}$) is also commonly used.

Speed and **velocity** are both quantities that give an indication of how quickly the position of an object is changing. Both terms are in common use and are often assumed to have the same meaning. In physics, however, these two terms have different definitions.

Instantaneous speed and velocity

Instantaneous speed and instantaneous velocity give a measure of how fast something is moving at a particular point in time. The speedometer on a car or bike indicates instantaneous speed.

If a speeding car is travelling north and is detected on a police radar gun at $150\,km\,h^{-1}$, it indicates that this car's instantaneous speed is $150\,km\,h^{-1}$, while its instantaneous velocity is $150\,km\,h^{-1}$ north. Notice that the instantaneous speed is equal to the magnitude of the instantaneous velocity. This is always the case for instantaneous speed and velocity.

Average speed and velocity

Average speed and *average velocity* both give an indication of how fast an object is moving over a time interval.

$$\text{average/mean speed } v_{av} = \frac{\text{distance travelled}}{\text{time taken}} = \frac{d}{t}$$

$$\text{average/mean velocity } v_{av} = \frac{\text{displacement}}{\text{time taken}} = \frac{s}{t} = \frac{v+u}{2}$$

A direction (such as north, south, up, down, left, right, positive, negative) must be given when describing a velocity. The direction will always be the same as that of the displacement.

Average or mean speed is equal to instantaneous speed only when a body is moving in uniform motion (that is, if it moves at a constant speed). The mean speed of a car that takes 30 minutes to travel 20 km from Perth to Sorrento is $40\,km\,h^{-1}$. However, this does not mean that the car travelled the whole distance at this speed. In fact it is more likely that the car was moving at $60\,km\,h^{-1}$ for a significant amount of time, while for some of the time the car would not be moving at all.

Similar to the relationship between distance and displacement, average speed will be equal to the magnitude of average velocity only if the body is moving in a straight line and does not change direction.

For example, in a race around a circular track like the velodrome shown in Figure 7.1.6, regardless of the average speed for a complete lap the magnitude of the average velocity will be zero, because the displacement is zero.

FIGURE 7.1.6 Anna Meares won the UCI world championship in 2013. She rode 500 m in a world record time of 32.836 s. Her average speed was $55.6\,km\,h^{-1}$ but her average velocity was zero.

EXTENSION

How police measure the speeds of cars

Road accidents cause the deaths of about 1200 people in Australia each year and many times this number are seriously injured. Numerous steps have been taken to reduce the number of road fatalities. Some of these include random alcohol and drug testing, speed cameras, mandatory wearing of bicycle helmets and the zero blood alcohol level for probationary drivers.

One of the main causes of road trauma is speeding. In their efforts to combat speeding motorists, police employ a variety of speed-measuring devices. One such device is shown in Figure 7.1.7.

FIGURE 7.1.7 Speed cameras on poles.

Speed camera radar

Camera radar units are usually placed in parked, unmarked vehicles. These units emit a radar signal frequency of 24.15 GHz (2.415×10^{10} Hz). The radar antenna has a parabolic reflector that enables the unit to produce a directional radar beam that is 5° wide, allowing individual vehicles to be targeted. The radar range and field of vision for a camera is shown in Figure 7.1.8. The radar signal allows speeds to be determined by the Doppler effect, where the reflected radar signal from an approaching vehicle has a higher frequency than the original signal. Similarly, the reflected signal from a receding vehicle has a lower frequency. This change in frequency or 'Doppler shift' is processed by the unit and gives a measurement of the instantaneous speed of the target vehicle.

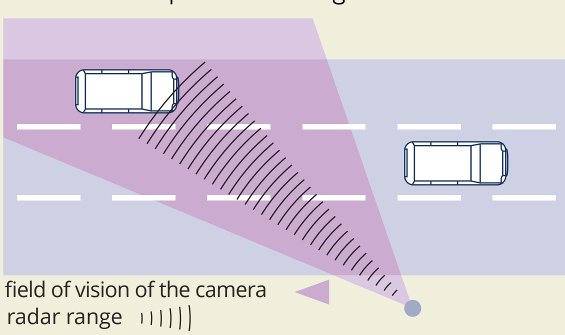


FIGURE 7.1.8 Diagram showing the visual range of a speed camera.

Camera radar units are capable of targeting a single vehicle up to 1.2 km away. In traffic, the units can distinguish between individual cars and take two photographs per second. The photographs and infringement notices are mailed to the offending motorists.

Laser speed guns

Speed guns are used by police to obtain an instant measure of the speed of an approaching or receding vehicle. The unit is usually handheld and is aimed directly at a vehicle using a target sight. It emits a pulse of infra-red radiation frequency of 331 THz (3.31×10^{14} Hz). As with camera radar units, the speed is determined by the Doppler shift produced by the target vehicle. The infra-red pulse is very narrow and directional, just 0.17° wide. This allows vehicles to be targeted with great precision. Handheld units can be used at distances up to 800 m. If the vehicle's speed registers over the limit, police are likely to pull the driver over.

Fixed speed cameras

Fixed speed cameras obtain their readings by using a system of three strips with piezoelectric sensors in them across the road (see Figure 7.1.9). The strips respond to the pressure as the car drives over them and create an electrical pulse that is detected. By knowing the distance between the strips and measuring the time that the car takes to travel across them, the speed of the car can be determined. This is actually measuring the average speed of the car, but by placing the strips close together the average speed gives a very good approximation of the instantaneous speed.

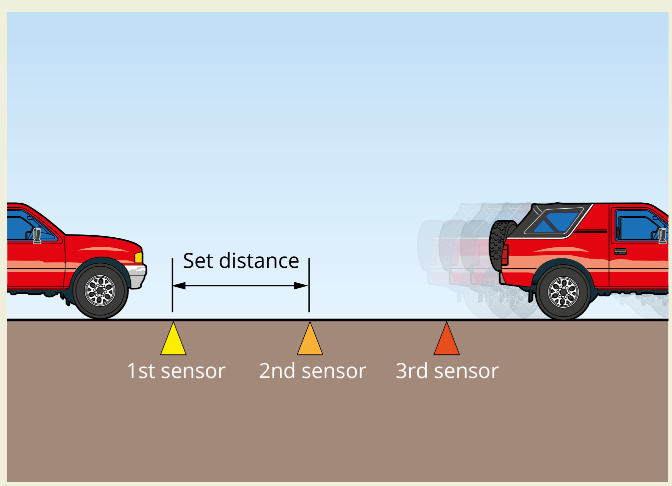


FIGURE 7.1.9 Fixed speed cameras record the speed of a car twice by measuring the time the car takes to travel over a series of three sensor strips embedded in the roadway.

Converting $km\,h^{-1}$ to $m\,s^{-1}$

You should be familiar with $100\,km\,h^{-1}$ as it is the speed limit for most freeways and country roads in Australia. Cars that maintain this speed would travel 100 km in 1 hour. Since there are 1000 metres in 1 kilometre and 3600 seconds in 1 hour ($60\,s \times 60\,min$), this is the same as travelling 100 000 m in 3600 s.

$$\begin{aligned} 100\,km\,h^{-1} &= 100 \times 1000\,m\,h^{-1} \\ &= 100\,000\,m\,h^{-1} \\ &= \frac{100\,000}{3600}\,m\,s^{-1} \\ &= 27.8\,m\,s^{-1} \end{aligned}$$

So $km\,h^{-1}$ can be converted to $m\,s^{-1}$ by multiplying by $\frac{1000}{3600}$ (or dividing by 3.6).

Converting $m\,s^{-1}$ to $km\,h^{-1}$

A champion Olympic sprinter can run at an average speed of close to $10\,m\,s^{-1}$. Each second, the athlete will travel approximately 10 metres. At this rate, in 1 hour the athlete would travel $10 \times 3600 = 36\,000\,m = 36\,km$.

$$\begin{aligned} 10\,m\,s^{-1} &= 10 \times 3600\,m\,h^{-1} \\ &= 36\,000\,m\,h^{-1} \\ &= \frac{36\,000}{1000}\,km\,h^{-1} \\ &= 36\,km\,h^{-1} \end{aligned}$$

So $m\,s^{-1}$ can be converted to $km\,h^{-1}$ by multiplying by $\frac{3600}{1000}$ or 3.6.

When converting a speed from one unit to another, it is important to think about the speeds to ensure that your answers make sense. The diagram in Figure 7.1.10 summarises the conversion between units for speed.

PHYSICSFILE

Reaction time

Drivers are often distracted by loud music or phone calls. These distractions result in many accidents and deaths on the road. If cars are travelling at high speeds, they will travel a large distance in the time that the driver takes just to apply the brakes. A short reaction time is very important for all road users. This is easy to understand given the relationship between speed, distance and time.

distance travelled = $v \times t$

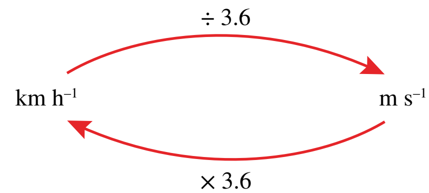


FIGURE 7.1.10 Rules for converting between $m\,s^{-1}$ and $km\,h^{-1}$.

PHYSICS IN ACTION

Alternative units for speed and distance

Metres per second is the standard unit for measuring speed because it is derived from the standard unit for distance (metres) and the standard unit for time (seconds). However, alternative units are often used to better suit a certain application.

The speed of a boat is usually measured in knots, where 1 knot = $0.51\,m\,s^{-1}$. This unit originated in the nineteenth century, when the speed of sailing ships would be measured by allowing a rope, with knots tied at regular intervals, to be dragged by the water through a sailor's hands. By counting the number of knots that passed through the sailor's hands, and measuring the time taken for this to happen, the average speed formula could be applied to estimate the speed of the ship.

The speed of very fast aeroplanes, such as the one in Figure 7.1.11, can be measured in Mach numbers. One Mach (referred to as Mach 1) is equal to the speed of sound, which is $340\,m\,s^{-1}$. Mach 2 is equal to $680\,m\,s^{-1}$, or twice the speed of sound.

The light-year is an alternative unit for measuring distance. The speed of light in a vacuum is nearly $300\,000\,km\,s^{-1}$. One light-year is the distance that light travels in one year. Because distances between objects in the universe are so large, astronomers use the light-year to measure distances. It takes over 4 years for light to travel from our nearest star (Alpha Centauri) to Earth. That means the distance from Earth to our nearest star is over 4 light-years. Light takes approximately 8.5 minutes to travel from the Sun to Earth, so it could be said that the Sun is 8.5 light-minutes away.

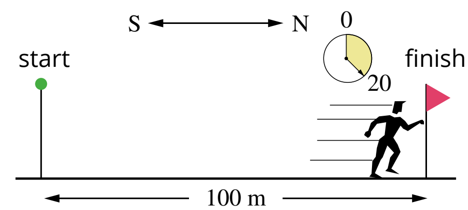
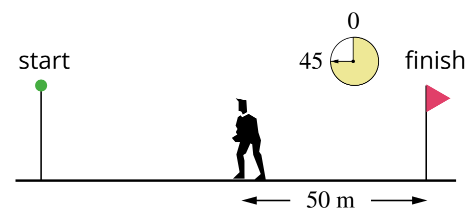
FIGURE 7.1.11 Modern fighter aeroplanes are able to fly at speeds above Mach 1.

Worked example 7.1.1

AVERAGE VELOCITY AND CONVERTING UNITS

Sam is an athlete performing a training routine by running back and forth along a straight stretch of running track. He jogs 100 m north in a time of 20 s, then turns and walks 50 m south in a further 25 s before stopping.

a What is Sam's average velocity in $m\,s^{-1}$?

Thinking	**Working**
Calculate the displacement. Remember that total displacement is the sum of individual displacements. Sam's total journey consists of two displacements: 100 north and then 50 m south. Take north to be the positive direction.	s = sum of displacements = 100 m north + 50 m south = 100 + (–50) = +50 m or 50 m north  
Work out the total time taken for the journey.	Time taken = 20 + 25 = 45 s
Substitute the values into the velocity equation.	Displacement, s, is 50 m north. Time taken, t, is 45 s. Average velocity $v_{av} = \frac{s}{t}$ $= \frac{50}{45}$ $= 1.1\ m\,s^{-1}$
Velocity is a vector, so a direction must be given.	$1.1\ m\,s^{-1}$ north

b What is the magnitude of Sam's average velocity in $km\,h^{-1}$?

Thinking	**Working**
Convert from $m\,s^{-1}$ to $km\,h^{-1}$ by multiplying by 3.6.	$v_{av} = 1.1\ m\,s^{-1}$ $= 1.1 \times 3.6$ $= 4.0\ km\,h^{-1}$ north
As the magnitude of the velocity is needed, the direction is not required in this answer.	Magnitude of $v_{av} = 4.0\ km\,h^{-1}$

c What is Sam's average speed in $m\,s^{-1}$?	
Thinking	**Working**
Calculate the distance. Remember that distance is the length of the path covered in the entire journey. The direction does not matter. Sam travels 100 m in one direction and then 50 m in the other direction.	$s = 100 + 50$ $= 150\,m$
Work out the total time taken for the journey.	$20 + 25 = 45\,s$
Substitute the values into the speed equation.	Distance, s, is 150 m. Time taken, t, is 45 s. Average speed $v_{av} = \frac{s}{t}$ $= \frac{150}{45}$ $= 3.3\,m\,s^{-1}$

d What is Sam's average speed in $km\,h^{-1}$?	
Thinking	**Working**
Convert from $m\,s^{-1}$ to $km\,h^{-1}$ by multiplying by 3.6.	Average speed $v_{av} = 3.3\,m\,s^{-1}$ $= 3.3 \times 3.6$ $= 12\,km\,h^{-1}$

Worked example: Try yourself 7.1.1

AVERAGE VELOCITY AND CONVERTING UNITS

Sally is an athlete performing a training routine by running back and forth along a straight stretch of running track. She jogs 100 m west in a time of 20 s, then turns and walks 160 m east in a further 45 s before stopping.

a What is Sally's average velocity in $m\,s^{-1}$?

b What is the magnitude of Sally's average velocity in $km\,h^{-1}$?

c What is Sally's average speed in $m\,s^{-1}$?

d What is Sally's average speed in $km\,h^{-1}$?

PHYSICSFILE

Breaking the speed limit

Over the past 100 years, advances in engineering and technology have led to the development of faster machines. Cars, planes and trains can now move people at speeds that were thought to be both unattainable and life-threatening a century ago.

The 1-mile land-speed record is $1220\,km\,h^{-1}$ ($339\,m\,s^{-1}$). This was set in 1997 in Nevada by Andy Green driving his jet-powered *Thrust* SSC.

The fastest combat jet is the MiG-25. In 1976 it reached a speed of $3800\,km\,h^{-1}$ ($1056\,m\,s^{-1}$), which is more than three times the speed of sound.

The fastest speed recorded by a train is $575\,km\,h^{-1}$ ($160\,m\,s^{-1}$) in 2007 by the French TGV *Atlantique*, although it does not reach this speed during normal operations.

In 2007, Markus Stoeckl of Austria set a new speed record for mountain biking. He reached a speed of $210\,km\,h^{-1}$ racing down a ski slope in Chile. He is pictured in Figure 7.1.12. This record was broken by Eric Barone in 2015 with a speed of $223.3\,km\,h^{-1}$.

FIGURE 7.1.12 Markus Stoeckl setting a new speed record for mountain biking in 2007.

7.1 Review

SUMMARY

- Position defines the location of an object with respect to a defined origin.
- Distance travelled, d, tells us how far an object has actually travelled. Distance travelled is a scalar.
- Displacement, s, is a vector and is defined as the change in position of an object in a given direction: s = final position – initial position.
- The average or mean speed of a body, v_{av}, is defined as the rate of change of distance and is a scalar quantity:
  average/mean speed $v_{av} = \frac{\text{distance travelled}}{\text{time taken}} = \frac{d}{t}$
- The average or mean velocity of a body, v_{av}, is defined as the rate of change of displacement and is a vector quantity:
  average/mean velocity
  $v_{av} = \frac{\text{displacement}}{\text{time taken}} = \frac{s}{t} = \frac{v + u}{2}$
- To convert from $m\,s^{-1}$ to $km\,h^{-1}$, multiply by 3.6.
- To convert from $km\,h^{-1}$ to $m\,s^{-1}$, divide by 3.6.
- The SI unit for both speed and velocity is metres per second ($m\,s^{-1}$).

KEY QUESTIONS

1 An student jogs one lap of the 400 m track in 2 minutes. Calculate

a her average speed

b her average velocity

2 A girl swims ten lengths of a 25 m pool. Which one or more of the following statements correctly describes her distance travelled and displacement?

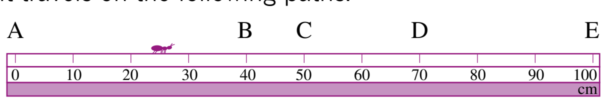
A Her distance travelled is zero.

B Her displacement is zero.

C Her distance travelled is 250 m.

D Her displacement is 250 m.

3 A confused ant is walking back and forth along a metre ruler, as show in the figure below. Taking the right as positive, determine both the size of the displacement and the distance travelled by the ant as it travels on the following paths.



a A to B

b C to B

c C to D

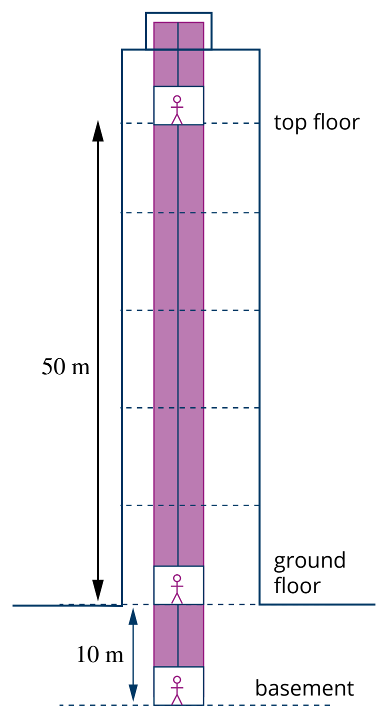
d C to E and then to D

4 During a training ride, a cyclist rides 50 km north then 30 km south.

a What is the distance travelled by the cyclist during the ride?

b What is the displacement of the cyclist for this ride?

5 A lift in a city building, shown in the figure below, carries a passenger from the ground floor down to the basement, then up to the top floor.



a What is the displacement of the lift as it travels from the ground floor to the basement?

b What is the displacement of the lift as it travels from the basement to the top floor?

c What is the distance travelled by the lift during this entire trip?

d What is the displacement of the lift during this entire trip?

6 A car travelling at a constant speed was timed over 400 m and was found to cover the distance in 12 s.

- **a** What was the car's average speed?
- **b** The driver was distracted and his reaction time was 0.75 s before applying the brakes. How far did the car travel in this time?

7 A cyclist travels 25 km in 90 minutes.

- **a** What is her average speed in $km\,h^{-1}$?
- **b** What is her average speed in $m\,s^{-1}$?

8 Liam pushes his toy truck 5 m east, then stops it and pushes it 4 m west. The entire motion takes 10 s.

- **a** What is the truck's average speed?
- **b** What is the truck's average velocity?

9 Mihi rides her bicycle to school and travels 2.5 km south in 15 min.

- **a** Calculate her average speed in kilometres per hour ($km\,h^{-1}$).
- **b** What was her average velocity in metres per second ($m\,s^{-1}$)?

10 An athlete in training for a marathon runs 10 km north along a straight road before realising that she has dropped her drink bottle. She turns around and runs back 3 km to find her bottle, then resumes running in the original direction. After running for 1.5 h, the athlete reaches 15 km from her starting position and stops.

- **a** What is the distance travelled by the athlete during the run?
- **b** What is the athlete's displacement during the run?
- **c** What is the average speed of the athlete in $km\,h^{-1}$?
- **d** What is the athlete's average velocity in $km\,h^{-1}$?

7.2 Acceleration

If you have been on a train as it pulled out of the station, you have experienced acceleration. If you have been in an aeroplane as it has taken off along a runway, you will have experienced a much greater acceleration. Astronauts and fighter pilots experience enormous accelerations that would make an untrained person lose consciousness. **Acceleration**, which is a measure of how quickly velocity changes, will be discussed in this section.

FINDING THE CHANGE IN VELOCITY AND SPEED

The velocity and speed of everyday objects are changing all the time. Examples of these are when a car moves away as the traffic lights turn green, when a tennis ball bounces or when you travel on a rollercoaster. If the initial and final velocity of an object are known, its change in velocity can be calculated.

To find the change, Δ, in any physical quantity, including speed and velocity, the initial value is taken away from the final value:

$$\Delta v = v - u$$

Change in speed is the final speed minus the initial speed:

$$\Delta v = v - u$$

where u is the initial speed (in m s^{-1})

v is the final speed (in m s^{-1})

Δv is the change in speed (in m s^{-1}).

Since speed is a scalar, direction is not required.

Change in velocity is the final velocity minus the initial velocity:

$$\Delta v = v - u$$

where u is the initial velocity (in m s^{-1})

v is the final velocity (in m s^{-1})

Δv is the change in velocity (in m s^{-1}).

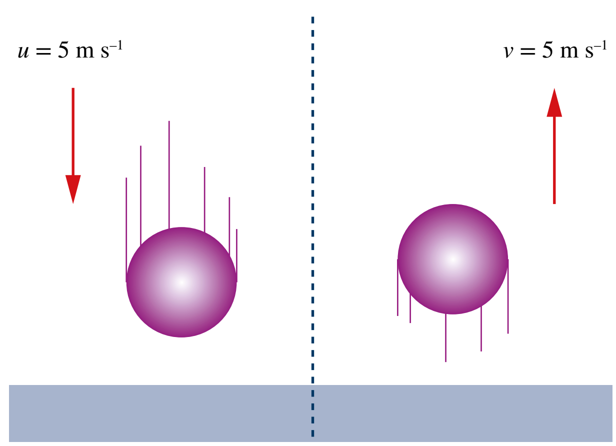
Since velocity is a vector, this should be done by performing a vector subtraction. As for all vectors, direction is required.

Vector subtraction was covered in detail in Section 6.2 on page 181.

Worked example 7.2.1

CHANGE IN SPEED AND VELOCITY PART 1

A golf ball is dropped onto a concrete floor and strikes the floor at 5.0 m s^{-1}. It then rebounds at 5.0 m s^{-1}.



a What is the change in speed of the ball?	
Thinking	**Working**
Find the values for the initial speed and the final speed of the ball.	$u = 5.0\,m\,s^{-1}$ $v = 5.0\,m\,s^{-1}$
Substitute the values into the change in speed equation: $\Delta v = v - u$	$\Delta v = v - u$ $= (5.0) - (5.0)$ $= 0.0\,m\,s^{-1}$

b What is the change in velocity of the ball?	
Thinking	**Working**
Velocity is a vector. Apply the sign convention to replace the directions.	$u = 5.0\,m\,s^{-1}$ down $= -5.0\,m\,s^{-1}$ $v = 5.0\,m\,s^{-1}$ up $= +5.0\,m\,s^{-1}$
As this is a vector subtraction, reverse the direction of u to get $-u$.	$u = -5.0\,m\,s^{-1}$, therefore $-u = +5.0\,m\,s^{-1}$
Substitute the values into the change in velocity equation: $\Delta v = v + (-u)$.	$\Delta v = v + (-u)$ $= (+5.0) + (+5.0)$ $= +10\,m\,s^{-1}$
Apply the sign convention to describe the direction.	$\Delta v = 1.0 \times 10^1\,m\,s^{-1}$ up

Worked example: Try yourself 7.2.1

CHANGE IN SPEED AND VELOCITY PART 1

A golf ball is dropped onto a concrete floor and strikes the floor at $9.0\,m\,s^{-1}$. It then rebounds at $7.0\,m\,s^{-1}$.

a What is the change in speed of the ball?

b What is the change in velocity of the ball?

ACCELERATION

Consider the following information about the velocity of a car that starts from rest as shown in Figure 7.2.1.

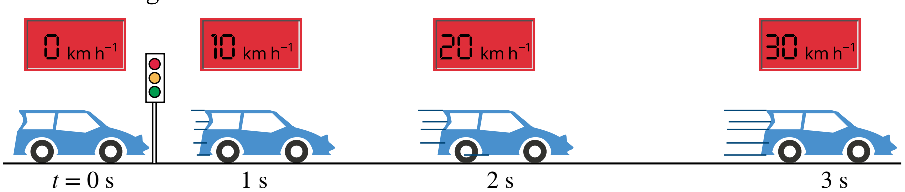


FIGURE 7.2.1 A car's acceleration as it increases in velocity from $0\,km\,h^{-1}$ to $30\,km\,h^{-1}$.

The velocity of the car pictured above increases by $10\,km\,h^{-1}$ each second. In other words, its velocity changes by $+10\,km\,h^{-1}$ per second. This is stated as an acceleration of +10 kilometres per hour per second or $+10\,km\,h^{-1}\,s^{-1}$. More commonly in physics, velocity information is given in metres per second.

The athlete in Figure 7.2.2 takes 3 seconds to come to a stop at the end of a race.

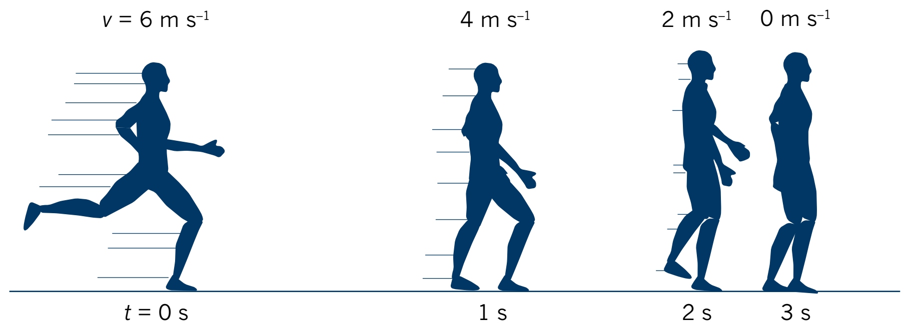


FIGURE 7.2.2 The velocity of the athlete changes by $-2\,\text{m}\,\text{s}^{-1}$ each second. The acceleration is $-2\,\text{m}\,\text{s}^{-2}$.

The velocity of the athlete changes by $-2\,\text{m}\,\text{s}^{-1}$ each second, so the acceleration is -2 metres per second per second. This is usually expressed as -2 metres per second squared or $-2\,\text{m}\,\text{s}^{-2}$.

A negative acceleration can mean that the object is slowing down in the direction of travel, as is the case with the athlete in Figure 7.2.2. A negative acceleration can also mean speeding up but in the opposite direction.

As acceleration is a vector quantity, vector diagrams can be used to calculate resultant accelerations of an object. Vector diagrams were covered in Chapter 6.

Average acceleration

As with speed and velocity, the average acceleration of an object can also be calculated.

Average acceleration, a_{av}, is the rate of change of velocity:

$$a = \frac{\text{change in velocity}}{\text{time taken}}$$
$$= \frac{v}{t}$$
$$= \frac{v - u}{t}$$

where a is the acceleration (in $\text{m}\,\text{s}^{-2}$)
v is the final velocity (in $\text{m}\,\text{s}^{-1}$)
u is the initial velocity (in $\text{m}\,\text{s}^{-1}$)
t is the time interval (s).

Worked example 7.2.2

CHANGE IN SPEED AND VELOCITY PART 2

A golf ball is dropped onto a concrete floor and strikes the floor at $5.0\,\text{m}\,\text{s}^{-1}$. It then rebounds at $5.0\,\text{m}\,\text{s}^{-1}$. The contact with the floor lasts for 25 ms.
What is the average acceleration of the ball during its contact with the floor?

Thinking	Working
Note the values you will need in order to find the average acceleration (initial velocity, final velocity and time). Convert ms into s by dividing by 1000. (Note that Δv was calculated for this situation in the previous Worked example.)	$u = -5.0\,\text{m}\,\text{s}^{-1}$ $-u = 5.0\,\text{m}\,\text{s}^{-1}$ $v = 5.0\,\text{m}\,\text{s}^{-1}$ $\Delta v = v - u = 10\,\text{m}\,\text{s}^{-1}$ up $t = 25\,\text{ms}$ $= 0.025\,\text{s}$

Substitute the values into the average acceleration equation.	$a = \frac{\text{change in velocity}}{\text{time taken}}$ $a = \frac{v}{t} = \frac{v - u}{t}$ $= \frac{10}{0.025}$ $= 400\,\text{m}\,\text{s}^{-2}$
Acceleration is a vector, so you must include a direction in your answer.	$a = 4.0 \times 10^2\ \text{m}\,\text{s}^{-2}$ up

Worked example: Try yourself 7.2.2

CHANGE IN SPEED AND VELOCITY PART 2

A golf ball is dropped onto a concrete floor and strikes the floor at $9.0\,\text{m}\,\text{s}^{-1}$. It then rebounds at $7.0\,\text{m}\,\text{s}^{-1}$. The contact time with the floor is 35 ms.

What is the average acceleration of the ball during its contact with the floor?

PHYSICSFILE

Human acceleration

In the 1950s, the United States Air Force used a rocket sled to determine the effect of extremely large accelerations on humans. One of these sleds is shown in Figure 7.2.3. The aim was to find out the greatest accelerations that humans could safely withstand in order to develop ejector seats for pilots.

The testing site consisted of an 800 m–long railway track and a sled with nine rocket motors. One volunteer, Colonel John Stapp, was strapped into the sled and accelerated to speeds of over $1000\,\text{km}\,\text{h}^{-1}$ in a very short time. Water scoops were used to stop the sled abruptly in just 0.35 s. This equates to a deceleration of greater than $400\,\text{m}\,\text{s}^{-2}$. The effects of these massive accelerations are evident on his face (Figure 7.2.4).

Colonel John Stapp was a human guinea pig who suffered a great deal of discomfort so that other pilots would benefit. Safer ejector seats and non-human crash test dummies were developed as a result of these experiments.

FIGURE 7.2.3 The rocket-powered sled used to test the effects of acceleration on humans.

FIGURE 7.2.4 Photos showing the distorted face of Colonel John Stapp.

7.2 Review

SUMMARY

- Change in speed is a scalar calculation:
  Δv = final speed – initial speed = $v - u$
- Change in velocity is a vector calculation:
  Δv = final velocity – initial velocity = $v - u$
- Acceleration is a vector. The acceleration of a body, a is defined as the rate of change of velocity:

$$a = \frac{\text{change in velocity}}{\text{time taken}}$$

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

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

- Acceleration is measured in metres per second per second ($\text{m}\,\text{s}^{-2}$).

KEY QUESTIONS

1 A radio-controlled car is travelling east at $10\,\text{km}\,\text{h}^{-1}$. It hits some sand and slows down to $3\,\text{km}\,\text{h}^{-1}$ east. Determine its change in speed.

2 A lump of Blu Tack is falling vertically at $5.0\,\text{m}\,\text{s}^{-1}$ and as it hits the floor it stops dead. Calculate its change in velocity during the collision.

3 A ping pong ball is falling vertically at $5.0\,\text{m}\,\text{s}^{-1}$. As it hits the floor, it rebounds at $3.0\,\text{m}\,\text{s}^{-1}$ up. Calculate its change in velocity during the bounce.

4 While playing soccer, Ashley is running north at $7.5\,\text{m}\,\text{s}^{-1}$. He slides along the ground and stops in 1.5 s. Calculate his average acceleration as he slides to a stop.

5 Olivia launches a model rocket vertically and it reaches a speed of $150\,\text{km}\,\text{h}^{-1}$ after 3.5 s. What is the magnitude of its average acceleration in $\text{km}\,\text{h}^{-1}\,\text{s}^{-1}$?

6 A squash ball travelling east at $25\,\text{m}\,\text{s}^{-1}$ strikes the front wall of the court and rebounds at $15\,\text{m}\,\text{s}^{-1}$ west. The contact time between the wall and the ball is 0.050 s. Use vector diagrams, where appropriate, to help you calculate:

a the change in speed of the ball.
b the change in velocity of the ball.
c the magnitude of the average acceleration of the ball during its contact with the wall.

7 A greyhound starts from rest and accelerates uniformly. Its velocity after 1.2 s is $8.0\,\text{m}\,\text{s}^{-1}$ south. Determine:

a the change in speed of the greyhound
b the change in velocity of the greyhound
c the magnitude of the acceleration of the greyhound.

8 How long does it take a vehicle travelling at $10.0\,\text{m}\,\text{s}^{-1}$ to reach $30.0\,\text{m}\,\text{s}^{-1}$ if it accelerates at $3.00\,\text{m}\,\text{s}^{-2}$?

9 A car travelling at $20.0\,\text{m}\,\text{s}^{-1}$ decelerates at $2.50\,\text{m}\,\text{s}^{-2}$. Calculate the time taken to stop.

10 A cyclist takes 4.00 s to slow down at $-3.00\,\text{m}\,\text{s}^{-2}$ and completely stop. Calculate the initial velocity of the cyclist.

7.3 Graphing position, velocity and acceleration over time

At times, even the motion of an object travelling in a straight line can be complicated. The object may travel forwards or backwards, speed up or slow down, or even stop. Where the motion remains in one dimension, the information can be presented in graphical form.

The main advantage of a graph compared with a table is that it allows the nature of the motion to be seen clearly. Information that is contained in a table is not as readily accessible or as easy to interpret as information presented graphically. This section examines position–time, velocity–time and acceleration–time graphs.

POSITION–TIME (x–t) GRAPHS

A position–time graph indicates the position, x, of an object at any time, t, for motion that occurs over an extended time interval. However, the graph can also provide additional information.

Consider Sophie, shown in Figure 7.3.1, swimming laps of a 50 m pool. Her position–time data are shown in Table 7.3.1. The starting point is treated as the origin for this motion.

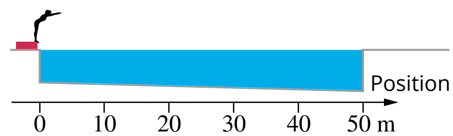


FIGURE 7.3.1 Swimmer standing at the end of a 50 m swimming pool.

TABLE 7.3.1 Positions and times of a swimmer completing 1.5 laps of a pool.

Time (s)	0	5	10	15	20	25	30	35	40	45	50	55	60
Position (m)	0	10	20	30	40	50	50	50	45	40	35	30	25

Analysis of Table 7.3.1 reveals several features of Sophie's swim. For the first 25 s, she swims at a constant rate. Every 5 s she travels 10 m in a positive direction, i.e. her velocity is $+2\,\text{m}\,\text{s}^{-1}$. Then, from 25 s to 35 s, her position does not change. She seems to be resting, as she is stationary for this 10 s interval. Finally, from 35 s to 60 s, she swims back towards the starting point, in a negative direction. On this return lap, she maintains a more leisurely rate of 5 m every 5 s, so her velocity is $-1\,\text{m}\,\text{s}^{-1}$. However, Sophie does not complete this lap but ends 25 m from the start. This data is shown more conveniently on the position–time graph in Figure 7.3.2.

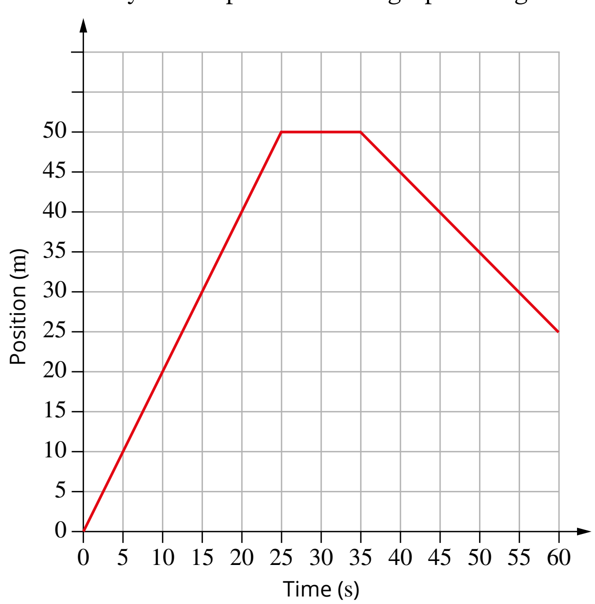


FIGURE 7.3.2 This position–time graph represents the motion of a swimmer travelling 50 m along a pool, then resting and swimming back towards the starting position. The swimmer finishes halfway along the pool.

The displacement, s, of the swimmer can be determined by comparing the initial and final positions. Her displacement between 20 s and 60 s is, for example:

$$\begin{aligned} s &= \text{final position} - \text{initial position} \\ &= 25 - 40 \\ &= -15\,\text{m} \end{aligned}$$

By further examining the graph, it can be seen that during the first 25 s, the swimmer has a displacement of +50 m. Thus her average velocity is $+2\,\text{m}\,\text{s}^{-1}$, i.e. $2\,\text{m}\,\text{s}^{-1}$ to the right, during this time. This value can also be obtained by finding the gradient of this section of the graph.

A straight-line in a position–time graph (Figure 7.3.3) indicates a uniform velocity. The slope (gradient) of the line is equal to the velocity of the object.

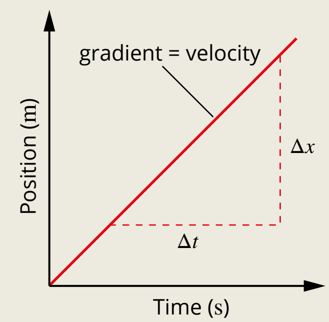


FIGURE 7.3.3 Position–time graph with gradient.

A positive velocity indicates that the object is moving in a positive direction and negative velocity indicates motion in a negative direction.

To confirm that the gradient of a position–time graph is a measure of velocity you can use **dimensional analysis**:

$$\text{Gradient of } x\text{–}t \text{ graph} = \frac{\text{rise}}{\text{run}} = \frac{x}{t}$$

The units of this gradient will be metres per second ($\text{m}\,\text{s}^{-1}$) so gradient is a measure of velocity. Note that the rise in the graph is the change in position, which is the definition of displacement; that is, $\Delta x = s$.

Non-uniform velocity

For motion with uniform (constant) velocity, the position–time graph will be a straight line, but if the velocity is non-uniform the graph will be curved. If the position–time graph is curved, the instantaneous velocity will be the gradient of the tangent to the line at the point of interest; the average velocity will be the gradient of the chord between two points. This is illustrated in Figure 7.3.4.

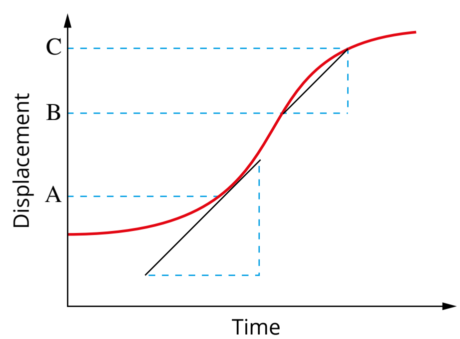
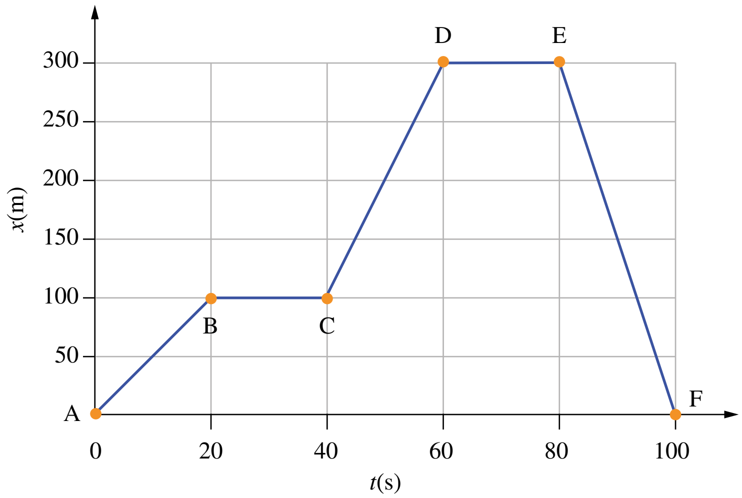


FIGURE 7.3.4 The instantaneous velocity at point A is the gradient of the tangent at that point. The average velocity between points B and C is the gradient of the chord between these points on the graph.

Worked example 7.3.1

ANALYSING A POSITION–TIME GRAPH

The motion of a cyclist is represented by the position–time graph below, with important features of the motion labelled A, B, C, D, E and F.



a What is the velocity of the cyclist between A and B?	
Thinking	**Working**
Determine the change in position (displacement) of the cyclist between A and B using: s = final position – initial position	At A, $x = 0$ m. At B, $x = 100$ m. $s = 100 - 0$ $= +100$ m or 100 m forwards (that is, away from the starting point)
Determine the time taken to travel from A to B.	$\Delta t = 20 - 0$ $= 20$ s
Calculate the gradient of the graph between A and B using: gradient of x–t graph $= \frac{\text{rise}}{\text{run}} = \frac{x}{t}$ Remember that $\Delta x = s$.	Gradient $= \frac{100}{20}$ $= 5$
State the velocity, using: gradient of x–t graph = velocity Velocity is a vector so a direction must be given.	Since the gradient is 5, the velocity is $+5\,\text{m s}^{-1}$ or $5\,\text{m s}^{-1}$ forwards.

b Describe the motion of the cyclist between B and C.	
Thinking	**Working**
Interpret the shape of the graph between B and C.	The graph is flat between B and C, indicating that the cyclist's position is not changing for this time. So the cyclist is not moving. If the cyclist is not moving, the velocity is $0\,\text{m s}^{-1}$.
You may confirm the result by calculating the gradient of the graph between B and C using: gradient of x–t graph $= \frac{\text{rise}}{\text{run}} = \frac{x}{t}$ Remember that $\Delta x = s$.	Gradient $= \frac{0}{20}$ $= 0$

Worked example: Try yourself 7.3.1

ANALYSING A POSITION–TIME GRAPH

Use the graph shown in Worked example 7.3.1 to answer the following questions.

a What is the velocity of the cyclist between E and F?
b Describe the motion of the cyclist between D and E.

VELOCITY–TIME (v–t) GRAPHS

Analysing motion

A graph of velocity, v, against time, t, shows how the velocity of an object changes with time. This type of graph is useful for analysing the motion of an object moving in a complex manner.

Consider the example of the girl in Figure 7.3.5. Aliyah is running back and forth along an aisle in a supermarket. A study of the velocity–time graph in Figure 7.3.5 reveals that Aliyah is moving with a positive velocity, i.e. in a positive direction, for the first 6 s. Between the 6 s mark and the 7 s mark she is stationary, then she runs in the reverse direction, i.e. has negative velocity, for the final 3 s.

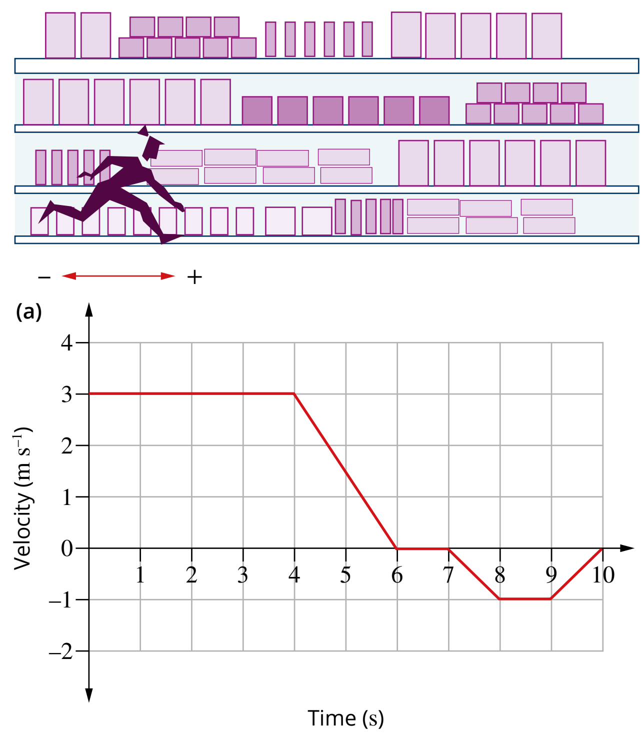


FIGURE 7.3.5 Diagram and v–t graph for a girl running along an aisle.

The graph shows Aliyah's velocity at each instant in time. She moves in a positive direction with a constant speed of 3 m s^{-1} for the first 4 s. From 4 s to 6 s, she continues moving in a positive direction but slows down. At 6 s, she comes to a stop for 1 s. During the final 3 s, she accelerates in the negative direction for 1 s then travels at a constant velocity of –1 m s^{-1} for 1 s. She then slows down and comes to a stop at 10 s. Remember that whenever the graph is below the time axis, velocity is negative, which indicates travel in the reverse direction. So she is travelling in the reverse direction for the last 3 s of her journey.

Finding displacement

A velocity–time graph can also be used to find the displacement of the object under consideration.

Displacement, s, is given by the area under a velocity–time graph, i.e. the area between the graph and the time axis. It is important to note that an area below the time axis indicates a negative displacement, i.e. motion in a negative direction.

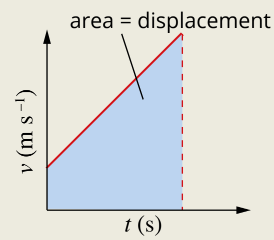


FIGURE 7.3.6 The area under a *v–t* graph gives displacement.

It is easier to see why the displacement is given by the area under the v–t graph when velocity is constant. For example, the graph in Figure 7.3.7 shows that in the first 6 s of motion, Aliyah moves with a constant velocity of $+3\,\text{m}\,\text{s}^{-1}$ for 4 s. Note that the area under the graph for this period of time is a rectangle. Her displacement, s, during this time can be determined by rearranging the formula for velocity:

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

$$\therefore s = v \times \Delta t$$

$$= \text{height} \times \text{base}$$

$$= \text{area under } v\text{–}t \text{ graph}$$

Aliyah then slows from $3\,\text{m}\,\text{s}^{-1}$ to zero in the next 2 s. To understand why the displacement for this period of time is given by the triangular area under the graph requires more complicated mathematics known as calculus, which is beyond the scope of this book.

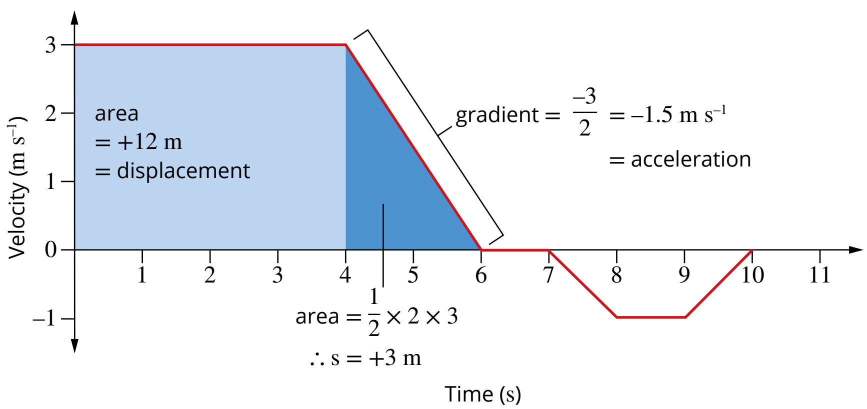


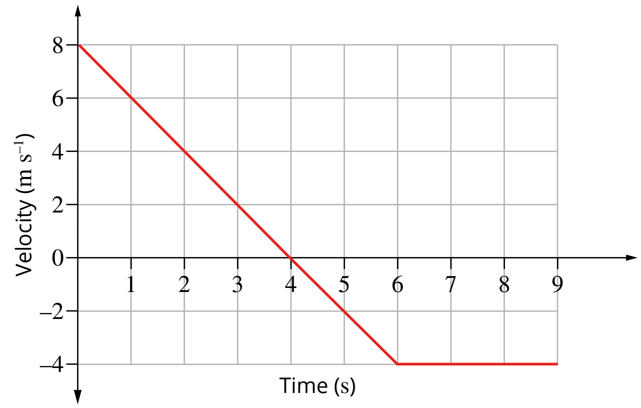
FIGURE 7.3.7 Area values as shown in a *v–t* graph.

From Figure 7.3.7, the area under the graph for the first 4 s gives Aliyah's displacement during this time, i.e. +12 m. The displacement from 4 s to 6 s is represented by the area of the darker blue triangle and is equal to +3 m. The total displacement during the first 6 s is +12 m + 3 m = +15 m.

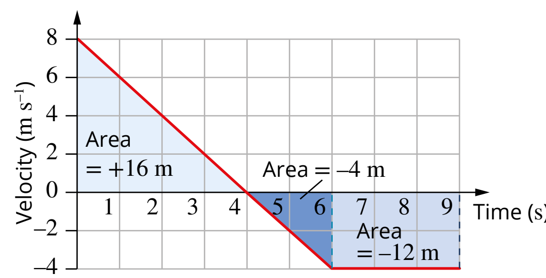
Worked example 7.3.2

ANALYSING A VELOCITY–TIME GRAPH

The motion of a radio-controlled car initially travelling east in a straight line across a driveway is represented by the graph below.



a What is the displacement of the car during the first 4 seconds?

Thinking	Working
Displacement is the area under the graph. So, calculate the area under the graph for the time period for which you want to find the displacement. Use displacement = $b \times h$ for squares and rectangles. Use displacement = $\frac{1}{2}(b \times h)$ for triangles.	 The area from 0 to 4 s is a triangle, so: area = $\frac{1}{2}(b \times h)$ $= \frac{1}{2} \times 4 \times 8$ $= +16\,\text{m}$
Displacement is a vector quantity, so a direction is needed.	displacement = 16 m east

b What is the average velocity of the car for the first 4 seconds?

Thinking	Working
Identify the equation and variables, and apply the sign convention.	$v = \frac{s}{t}$ $s = +16\,\text{m}$ $\Delta t = 4\,\text{s}$
Substitute values into the equation: $v = \frac{s}{t}$	$v = \frac{s}{t}$ $= \frac{+16}{4}$ $= +4\,\text{m}\,\text{s}^{-1}$
Velocity is a vector quantity, so a direction is needed.	$v_{av} = 4\,\text{m}\,\text{s}^{-1}$ east

Worked example: Try yourself 7.3.2

ANALYSING A VELOCITY–TIME GRAPH

Use the graph shown in Worked example 7.3.2 to answer the following questions.

a What is the displacement of the car from 4 to 6 seconds?

b What is the average velocity of the car from 4 to 6 seconds?

ACCELERATION FROM A VELOCITY–TIME (v–t) GRAPH

The acceleration of an object can also be found from a velocity–time graph.

The gradient of a velocity–time graph gives the average acceleration of the object over the time interval.

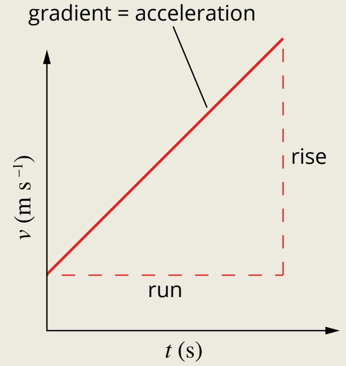


FIGURE 7.3.8 Gradient as displayed in a v–t graph.

Consider the motion of Aliyah in the 2 s interval between 4 s and 6 s on the graph in Figure 7.3.9. She is moving in a positive direction but slowing down from 3 m s^{-1} to rest.

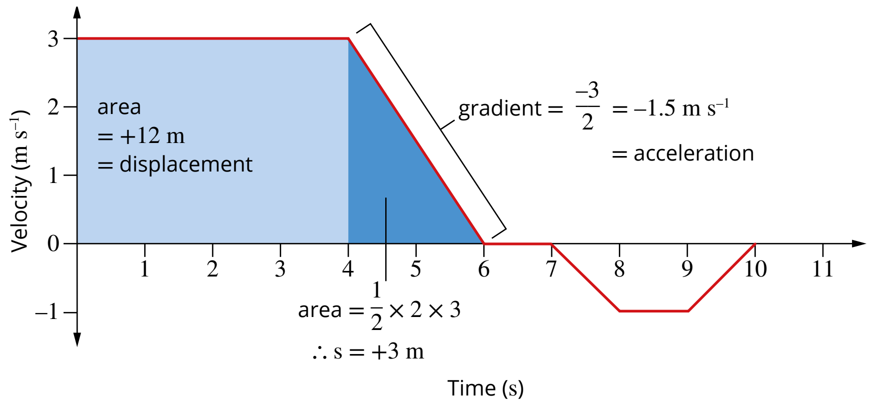


FIGURE 7.3.9 Acceleration as displayed in a v–t graph.

Her acceleration is:

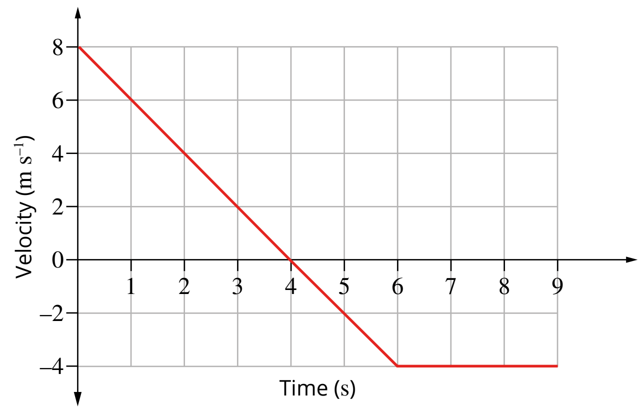
$$a = \frac{v}{t} = \frac{v-u}{t} = \frac{0-3}{2} = -1.5\,\text{m s}^{-2}$$

Since acceleration is the velocity change divided by time taken, it is given by the gradient of the v–t graph. As can be seen from Figure 7.3.9, the gradient of the line between 4 s and 6 s is -1.5 m s^{-2}.

Worked example 7.3.3

FINDING ACCELERATION USING A VELOCITY–TIME GRAPH

Consider the motion of the radio-controlled car described in Worked Example 7.2.3 initially travelling east in a straight line across a driveway as shown by the graph below.



What is the acceleration of the car during the first 4 s?

Thinking	Working
Acceleration is the gradient of a *v–t* graph. Calculate the gradient using: $\text{gradient} = \frac{\text{rise}}{\text{run}}$	$\text{Gradient from 0 to 4} = \frac{\text{rise}}{\text{run}}$ $= \frac{\delta v}{\delta t}$ $= \frac{0-8}{4-0}$ $= -2\,\text{m}\,\text{s}^{-2}$
Acceleration is a vector quantity, so a direction is needed. Note: In this case, the car is moving in the easterly direction and slowing down.	Acceleration = $-2\,\text{m}\,\text{s}^{-2}$ east (or $2\,\text{m}\,\text{s}^{-2}$ west)

Worked example: Try yourself 7.3.3

FINDING ACCELERATION USING A VELOCITY–TIME GRAPH

Use the graph shown in Worked example 7.3.3 to answer the following question.

What is the acceleration of the car during the period from 4 to 6 seconds?

DISTANCE TRAVELLED

A velocity–time graph can also be used to calculate the distance travelled by a moving object. The process of determining distance requires you to calculate the area under the v–t graph, similar to when calculating displacement. However, since distance travelled by an object always increases as the object moves, regardless of direction, you must add up all the areas between the graph and the time axis, regardless of whether the area is above or below the axis.

For example, Figure 7.3.10 shows the velocity–time graph of the radio-controlled car from Worked example 7.3.3. The area above the time-axis, which corresponds to motion in the positive direction, is +16 m, while the area below the axis, which corresponds to negative motion, consists of −4 m and −12 m. To calculate the total displacement, you would add up each displacement:

$$\begin{aligned}\text{total displacement} &= 16 + (-4) + (-12)\\ &= 16 - 16\\ &= 0\,\text{m}\end{aligned}$$

To calculate the total distance, you would add up the magnitude of the areas, by ignoring whether they are positive or negative:

$$\begin{aligned}\text{total distance} &= 16 + 4 + 12\\ &= 32\,\text{m}\end{aligned}$$

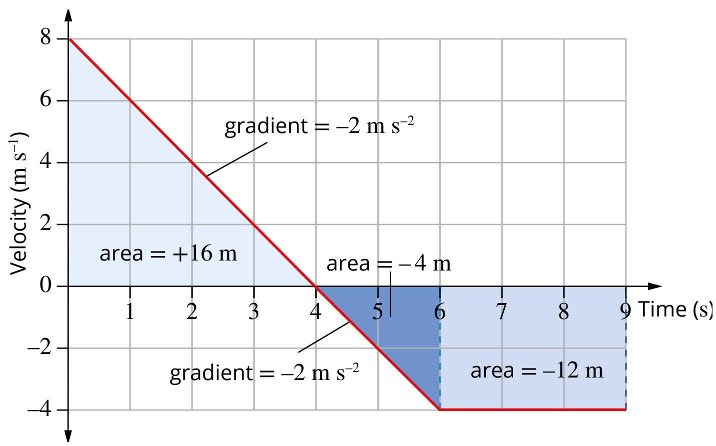


FIGURE 7.3.10 Both distance and displacement can be calculated by using the areas under the velocity–time graph.

Non-uniform acceleration

For motion with uniform (constant) acceleration, the velocity–time graph will be a straight line. For non-uniform acceleration the velocity–time graph will be curved. If the velocity–time graph is curved, the instantaneous acceleration will be the gradient of the tangent to the line at the point of interest; the average acceleration will be the gradient of the chord between two points. The displacement can still be calculated by finding the area under the graph, however you will need to make some estimations.

PHYSICSFILE

Area under graphs

The calculation of the area under a graph is useful in many areas of Physics.

Some examples include:

- power–time graph where the area represents the energy used over that period of time.
- force–time graph where the area represents the impulse or change in momentum over a period time (see Section 8.2)
- force–displacement graph where the area represents the work done or energy transferred while the forces are acting

ACCELERATION–TIME (a–t) GRAPHS

An acceleration–time graph simply indicates the acceleration of the object as a function of time. The area under an acceleration–time graph is found by multiplying an acceleration, a, and a period of time, Δt, value. The area gives a change in velocity, Δv, value:

$$\text{area} = a \times \Delta t = \Delta v$$

In order to establish the actual velocity of the object, its initial velocity must be known. Figure 7.3.11 shows both Aliyah's velocity versus time (v–t) and acceleration versus time (a–t) graphs.

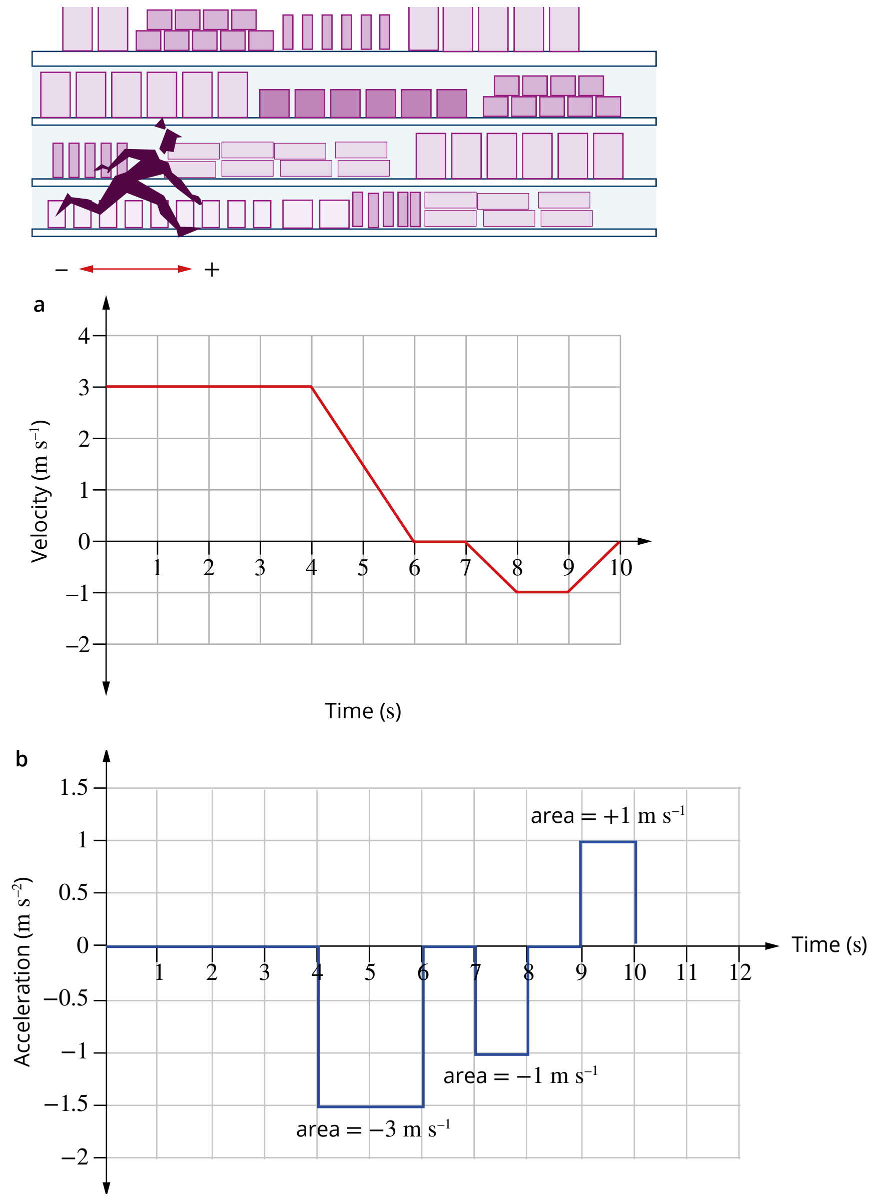


FIGURE 7.3.11 (a) Aliyah's velocity versus time (v–t) graph. (b) Aliyah's acceleration versus time (a–t) graph.

From 4 to 6 s, the area shows a Δv of $-3\,\text{m s}^{-1}$. This indicates that she has slowed down by $3\,\text{m s}^{-1}$ during this time. Her v–t graph confirms this fact. Her initial speed is $3\,\text{m s}^{-1}$, so she must be stationary ($v = 0$) after 6 s. This calculation could not be made without knowing her initial velocity.

EXTENSION

Graphing in Physics

Graphs in physics can be useful in solving problems as an alternative to using equations. An advantage of graphs is that they give a quick and easy picture of the relationship between the data that has been measured.

In this course, you will mainly analyse straight-line or linear graphs. A higher level of mathematical skills is required to analyse nonlinear graphs or curves.

Differentiation is used to find the gradients of curves, where the intervals of change are infinitely small.

For a straight line, gradient = $\frac{\text{rise}}{\text{run}} = \frac{x}{t}$

If you consider a curve to be a series of infinitesimally small straight lines, then Δx and Δt are also extremely small. The gradient then becomes $\frac{\delta x}{\delta t}$, where the lower-case Greek letter delta, δ, denotes an infinitesimally small change. This is often written as $\frac{dx}{dt}$, i.e. the derivative of x with respect to t.

For a velocity–time graph, the gradient gives the acceleration—i.e. acceleration = $\frac{dv}{dt}$.

You have already seen that displacement can be found by calculating the area under a velocity–time graph, for example by breaking up the area under the graph into rectangles and triangles. Similarly, taking extremely small sections and adding them all together again gives the area under a non-linear graph, as shown in Figure 7.3.12. This is called integration, and for a curve given by a function $f(x)$ can be written as:

area = $\int_a^b f(x)\,dx$

So for a velocity–time graph, finding the area under the graph to give the displacement:

$s = \int_{t1}^{t2} v\,dt$

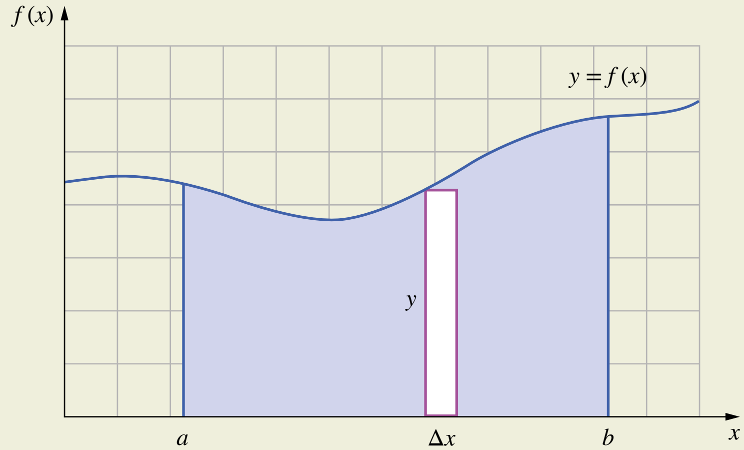


FIGURE 7.3.12 The small areas are added together by integration.

7.3 Review

SUMMARY

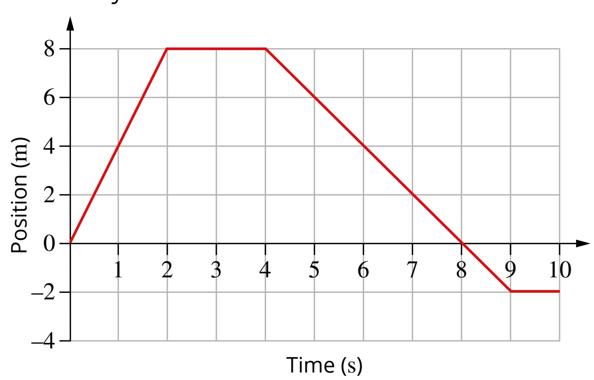
- A position–time graph can be used to determine the location of an object at any given time. Additional information can also be derived from the graph:
  - Displacement is given by the change in position of an object.
  - The velocity of an object is given by the gradient of the position–time graph.
  - If the position–time graph is curved, the gradient of the tangent at a point gives the instantaneous velocity.
- The gradient of a velocity–time graph is the acceleration of the object.
- The area under a velocity–time graph is the displacement of the object.
- The area under an acceleration–time graph is the change in velocity of the object.

KEY QUESTIONS

1 Which of the following does the gradient of a position–time graph represent?

A displacement
B acceleration
C time
D velocity

The following information relates to questions 2–6. The graph represents the straight-line motion of a radio-controlled toy car.



2 Describe the motion of the car in terms of its position.

3 What was the position of the toy car after:

a 2 s?
b 4 s?
c 6 s?
d 10 s?

4 When did the car return to its starting point?

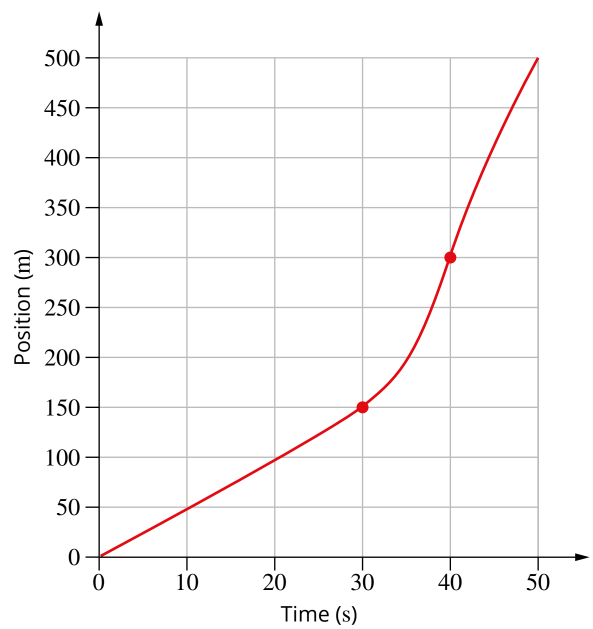
5 What was the velocity of the toy car:

a during the first 2 s?
b at 3 s?
c from 4 s to 8 s?
d at 8 s?
e from 8 s to 9 s?

6 During its 10 s motion, what was the car's:

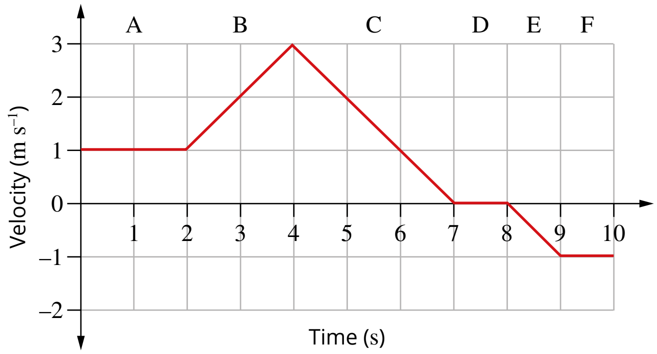
a distance travelled?
b displacement?

7 This position–time graph for a cyclist travelling north along a straight road is shown. Calculate the following information about the cyclist's motion.



a What was the average speed of the cyclist during the first 30 s?
b What was the average velocity of the cyclist during the final 10 s?
c What was the average velocity of the cyclist for the whole trip?

8 The graph in the figure below shows the motion of a dog running along a footpath.



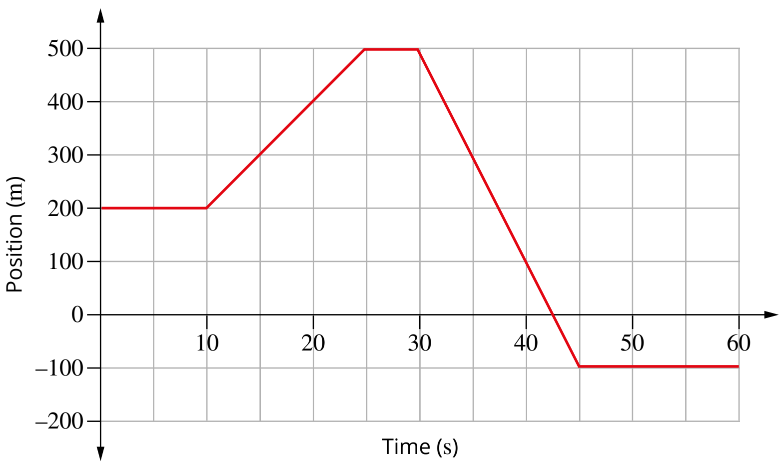
a What is the magnitude of the acceleration of the dog at $t = 1$ s?

b What is the magnitude of the acceleration of the dog at $t = 5$ s?

c What is the magnitude of the displacement of the dog for the first 7 s?

d What is the magnitude of the average velocity of the dog over the first 7 s?

9 The graph shows the position of a motorbike along a straight stretch of road as a function of time. The motorcyclist starts 200 m north of an intersection.

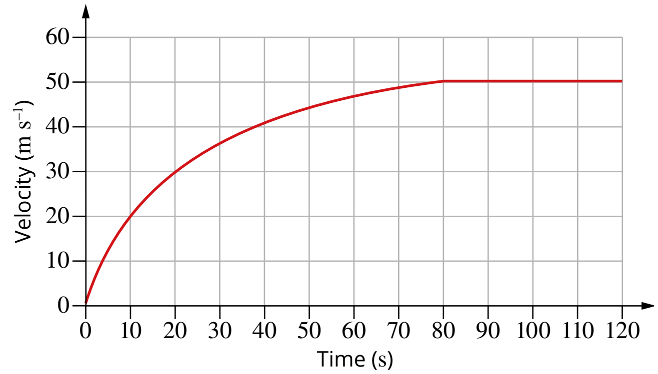


Calculate the instantaneous velocity of the motorcyclist at each of the following times:

a 15 s

b 35 s.

10 The straight-line motion of a high-speed intercity train is shown in the graph below.



a How long does it take the train to reach its cruising speed?

b What is the acceleration of the train 10 s after starting?

c What is the acceleration of the train 40 s after starting?

d By counting squares, or by another suitable method, approximate the displacement (in km) of the train after 120 s.

7.4 Equations for uniform acceleration

A graph is an excellent way of representing motion because it provides a great deal of information that is easy to interpret. However, a graph is time-consuming to draw and sometimes values have to be estimated rather than precisely calculated.

In the previous section, graphs of motion were used to determine quantities such as displacement, velocity and acceleration. This section examines a more powerful and precise method of solving problems involving *constant* or *uniform acceleration*. This method involves the use of a series of equations that can be derived from the basic definitions developed earlier.

DERIVING THE EQUATIONS

Consider an object moving in a straight line with an initial velocity, u, and a uniform acceleration, a, for a time interval, Δt. As u, v and a are vectors, and the motion is limited to one dimension, the sign and direction convention of right as positive and left as negative can be used. After a period of time, Δt, the object has changed its velocity from an initial velocity of u and is now travelling with a final velocity of v. Its acceleration will be given by:

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

If the initial time is 0 s, and the final time is t s, then $\Delta t = t$. The above equation can then be rearranged as:

$$v = u + at \quad \text{(i)}$$

The average velocity of the object is:

$$\text{average velocity } v_{av} = \frac{\text{displacement}}{\text{time taken}} = \frac{s}{t}$$

When acceleration is uniform, average velocity, v_{av}, can also be found as the average of the initial and final velocities:

$$v_{av} = \frac{1}{2}(u+v)$$

This relationship is shown graphically in Figure 7.4.1.

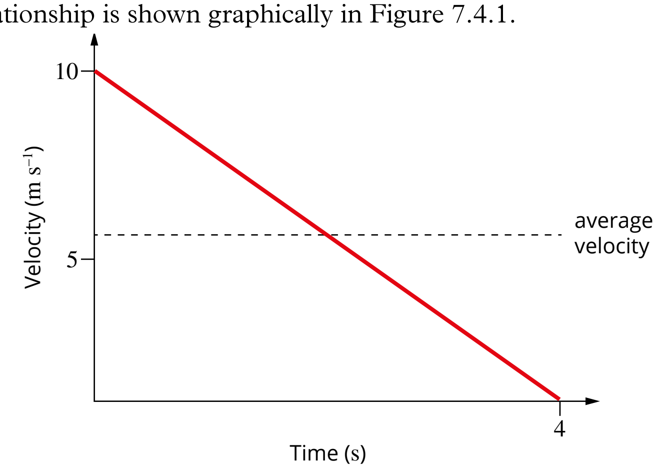


FIGURE 7.4.1 Uniform acceleration as displayed by a v–t graph.

So:

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

This gives:

$$s = \frac{1}{2}(u+v)t \quad \text{(ii)}$$

A graph describing constant acceleration motion is shown in Figure 7.4.2. For constant acceleration, the velocity is increasing by the same amount in each time interval, so the gradient of the v–t graph is constant.

The displacement, s, of the body is given by the area under the velocity–time graph. The area under the velocity–time graph, as shown in Figure 7.4.2, is given by the combined area of the rectangle and the triangle:

$$\text{Area} = s = ut + \frac{1}{2}(v - u) \times t$$

$$\text{As: } a = \frac{v - u}{t}$$

then: $v - u = at$, and this can be substituted for $v - u$:

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

$$s = ut + \frac{1}{2} \times at^2 \qquad \text{(iii)}$$

Making u the subject of equation (i) gives:

$$u = v - at$$

You might like to derive another equation yourself by substituting this into equation (ii). You will get:

$$s = vt - \frac{1}{2} \times at^2 \qquad \text{(iv)}$$

Rewriting equation (i) with t as the subject gives:

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

Now, if this is substituted into equation (ii):

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

$$= \frac{u + v}{2} \times \frac{v - u}{a}$$

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

Finally, transposing this gives:

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

Equations (i)–(v) are commonly used to solve problems in which acceleration is constant. They are summarised below.

$$v = u + at$$

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

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

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

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

where s is the displacement (in m)

u is the initial velocity (in m s^{-1})

v is the final velocity (in m s^{-1})

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

t is the time taken (in s).

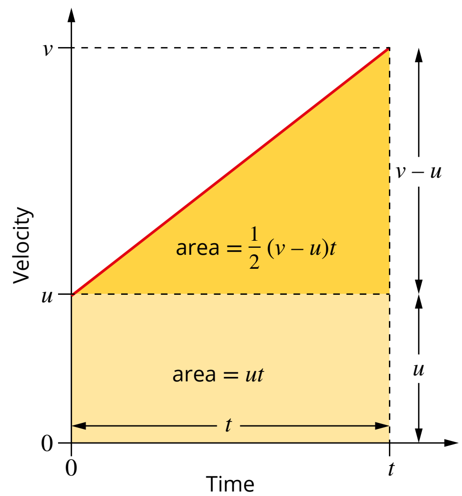


FIGURE 7.4.2 The area under a v–t graph broken up into a rectangle and a triangle.

SOLVING PROBLEMS USING EQUATIONS

When solving problems using these equations, it is important to think about the problem and try to visualise what is happening. Follow the steps below.

Step 1 Draw a simple diagram of the situation.

Step 2 Write down the information that has been given in the question. You might like to use the word 'suvat' as a memory trick to help you remember to list the variables in the order s, u, v, a and t. Use a sign convention to assign positive and negative values to indicate directions. Convert all units to SI form.

Step 3 Select the equation that matches your data. It should include three values that you know, and the one value that you want to solve.

Step 4 Use the appropriate number of significant figures in your answer.

Step 5 Include units with the answer and specify a direction if the quantity is a vector.

Worked example 7.4.1

USING THE EQUATIONS OF MOTION

A snowboarder in a race is travelling $10\,\text{m s}^{-1}$ north as he crosses the finishing line. He then decelerates uniformly, coming to a stop over a distance of $20\,\text{m}$.

a What is his acceleration as he comes to a stop?

Thinking	Working
Write down the known quantities as well as the quantity you are finding. Apply the sign convention that north is positive and south is negative.	Take all the information that you can from the question: • constant acceleration, so use equations for uniform acceleration • 'coming to a stop' means that the final velocity is zero. $s = +20\,\text{m}$ $u = +10\,\text{m s}^{-1}$ $v = 0\,\text{m s}^{-1}$ $a = ?$
Identify the correct equation to use.	$v^2 = u^2 + 2as$
Substitute known values into the equation and solve for a. Include units with the answer.	$v^2 = u^2 + 2as$ $0^2 = 10^2 + 2 \times a \times 20$ $0 = 100 + 40a$ $-100 = 40a$ $a = \frac{-100}{40}$ $= -2.5\,\text{m s}^{-2}$
Use the sign convention to state the answer with its direction.	$a = 2.5\,\text{m s}^{-2}$ south

b How long does he take to come to a stop?	
Thinking	**Working**
Write down the known quantities as well as the quantity you are finding. Apply the sign convention that north is positive and south is negative.	Take all the information that you can from the question: • constant acceleration, so use equations for uniform acceleration • 'coming to a stop' means that the final velocity is zero. $s = +20\,m$ $u = +10\,m\,s^{-1}$ $v = 0\,m\,s^{-1}$ $a = -2.5\,m\,s^{-2}$ $t = ?$
Identify the correct equation to use. Since you now know four values, any equation involving t will work.	$v = u + at$
Substitute known values into the equation and solve for t. Include units with the answer.	$v = u + at$ $0 = 10 + (-2.5) \times t$ $-10 = -2.5t$ $t = \frac{-10}{-2.5} = 4.0\,s$

c What is the average velocity of the snowboarder as he comes to a stop?	
Thinking	**Working**
Write down the known quantities as well as the quantity that you are finding. Apply the sign convention that north is positive and south is negative.	Take all the information that you can from the question: • constant acceleration, so we only need to find the average of the final and initial speeds. $u = +10\,m\,s^{-1}$ $v = 0\,m\,s^{-1}$ $v_{av} = ?$
Identify the correct equation to use.	$v_{av} = \frac{1}{2}(u + v)$
Substitute known values into the equation and solve for v_{av}. Include units with the answer.	$v_{av} = \frac{1}{2}(u + v)$ $= \frac{1}{2}(0 + 10)$ $= 5.0\,m\,s^{-1}$
Use the sign convention to state the answer with its direction.	$v_{av} = 5.0\,m\,s^{-1}$ north

Worked example: Try yourself 7.4.1

USING THE EQUATIONS OF MOTION

A snowboarder in a race is travelling $15\,m\,s^{-1}$ east as she crosses the finishing line. She then decelerates uniformly until coming to a stop over a distance of 30 m.

a What is her acceleration as she comes to a stop?

b How long does she take to come to a stop?

c What is the average velocity of the snowboarder as she comes to a stop?

7.4 Review

SUMMARY

- The following equations can be used for situations where there is a constant acceleration, where:
  - s = displacement (m)
  - u = initial velocity (m s^{-1})
  - v = final velocity (m s^{-1})
  - a = acceleration (m s^{-2})
  - t = time (s).
- $v = u + at$
- $s = \frac{1}{2}(u + v)t$
- $s = ut + \frac{1}{2}at^2$
- $s = vt + \frac{1}{2}at^2$
- $v^2 = u^2 + 2as$
- $v_{av} = \frac{s}{t} = \frac{u + v}{2}$
- A sign and direction convention for motion in one dimension needs to be used with these equations.

KEY QUESTIONS

1 A cyclist has a uniform acceleration as he rolls down a hill. His initial speed is 5 m s^{-1}, he travels a distance of 30 m and his final speed is 18 m s^{-1}. Which equation should be used to determine his acceleration?

A $v = u + at$

B $s = \frac{1}{2}(u + v)t$

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

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

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

2 A new-model Subaru travels with a uniform acceleration on a racetrack. It starts from rest and covers 400 m in 16 s.

a What is the magnitude of its average acceleration during this time?

b What is the final speed of the car in m s^{-1}?

c What is the car's final speed in km h^{-1}?

3 A Prius hybrid car starts from rest and accelerates uniformly in a positive direction for 8.0 s. It reaches a final speed of 16 m s^{-1}.

a What is the magnitude of the acceleration of the Prius?

b What is the magnitude of the average velocity of the Prius?

c What is the distance travelled by the Prius?

4 During its launch phase, a space rocket accelerates uniformly from rest to 160 m s^{-1} upwards in 4.0 s, then travels with a constant speed of 160 m s^{-1} for the next 5.0 s.

a What is the initial acceleration of the rocket?

b How far (in km) does the rocket travel in this 9.0 s period?

c What is the final speed of the rocket in km h^{-1}?

d What is the average speed of the rocket during the first 4.0 s?

e What is the average speed of the rocket during the 9.0 s motion?

5 While overtaking another cyclist, Ben increases his speed uniformly from 4.2 m s^{-1} to 6.7 m s^{-1} east over a time interval of 0.50 s.

a What is the magnitude of Ben's average acceleration during this time?

b How far does Ben travel while overtaking?

c What is Ben's average speed during this time?

6 A stone is dropped vertically into a lake. Which one of the following statements best describes the motion of the stone at the instant it enters the water?

A Its velocity and acceleration are both downwards.

B It has an upwards velocity and a net downwards acceleration.

C Its velocity and acceleration are both upwards.

D It has a downwards velocity and a net upwards acceleration.

7 A diver plunges headfirst into a diving pool while travelling at 28 m s^{-1} downwards. The diver enters the water and stops after a distance of 4.0 m. Consider the diver to be a single point located at her centre of mass and assume her acceleration through the water to be uniform.

a What is the magnitude of the average acceleration of the diver as she travels through the water?

b How long does the diver take to come to a stop?

c What is the speed of the diver after she has dived through 2.0 m of water?

8 A car is travelling along a straight road at 75 km h^{-1} east. In an attempt to avoid an accident, the motorist has to brake suddenly and stop the car.

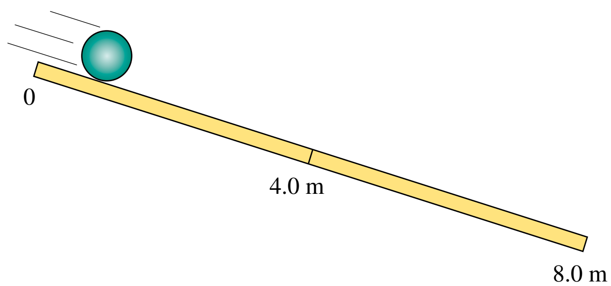
a What is the car's initial speed in m s^{-1}?

b If the reaction time of the motorist is 0.25 s, what distance does the car travel while the driver is reacting to apply the brakes?

c Once the brakes are applied, the car has an acceleration of –6.0 m s^{-2}. How far does the car travel while pulling up?

d What total distance does the car travel from the time the driver first notices the danger to when the car comes to a stop?

9 A billiard ball rolls from rest down a smooth ramp that is 8.0 m long. The acceleration of the ball is constant at 2.0 m s^{-2}.



a What is the speed of the ball when it is halfway down the ramp?

b What is the final speed of the ball?

c How long does the ball take to roll the first 4.0 m?

d How long does the ball take to travel the final 4.0 m?

10 A cyclist, Anna, is travelling at a constant speed of 12 m s^{-1} when she passes a stationary bus. The bus starts moving just as Anna passes, and it accelerates uniformly at 1.5 m s^{-2}.

a When does the bus reach the same speed as Anna?

b How long does the bus take to catch Anna?

c What distance has Anna travelled before the bus catches up?

7.5 Vertical motion

FIGURE 7.5.1 A stroboscopic image of a free-falling apple. The time elapsed between each image of the apple is the same but the distance it travels increases, which shows the apple is accelerating. Without air resistance, this rate of acceleration is the same for all objects.

Until 500 years ago, it was widely believed that the heavier an object was, the faster it would fall. This was the theory of Aristotle, and it lasted for 2000 years until the end of the Middle Ages. In the seventeenth century, the Italian scientist Galileo conducted experiments that showed that the mass of the object did not affect the rate at which it fell, as long as **air resistance** was not a factor.

It is now known that falling objects speed up because of gravity. Many people still think that heavier objects fall faster than light objects. This is not the case, but the confusion arises because of the effects of air resistance. This section examines the motion of falling objects.

ANALYSING VERTICAL MOTION

Some falling objects are affected by air resistance to a large extent, for example, feathers and balloons. This is why these objects do not speed up much as they fall. However, if air resistance can be ignored, all bodies in **free-fall** near the Earth's surface will move with an equal downwards acceleration. The stroboscopic image in Figure 7.5.1 clearly shows an apple accelerating as it falls, since the distance travelled by the apple between each photograph increases. In a vacuum, this rate of acceleration would be the same for a feather, a bowling ball, or any other object. The mass of the object does not matter.

At the Earth's surface, the acceleration due to gravity, g, is $9.8\,\mathrm{m\,s^{-2}}$ down and does not depend on whether the body has been thrown upwards or is falling down.

As an example, a coin that is dropped from rest will be moving at $9.8\,\mathrm{m\,s^{-1}}$ after 1 s, $19.6\,\mathrm{m\,s^{-1}}$ after 2 s, and so on. Each second its speed increases by $9.8\,\mathrm{m\,s^{-1}}$. The motion of a falling coin is illustrated in Figure 7.5.2.

However, if the coin was launched straight up at $19.6\,\mathrm{m\,s^{-1}}$, then after 1 s its speed would be $9.8\,\mathrm{m\,s^{-1}}$, and after 2 s it would be stationary. In other words, each second it would slow down by $9.8\,\mathrm{m\,s^{-1}}$. The motion of a coin thrown vertically upwards is shown in Figure 7.5.3.

So, regardless of whether the coin is falling or is tossed, its speed changes at the same rate. The speed of the falling coin *increases* by $9.8\,\mathrm{m\,s^{-1}}$ each second and the speed of the rising coin *decreases* by $9.8\,\mathrm{m\,s^{-1}}$ each second. That means that the acceleration of the coin due to gravity is $9.8\,\mathrm{m\,s^{-2}}$ downwards in both cases.

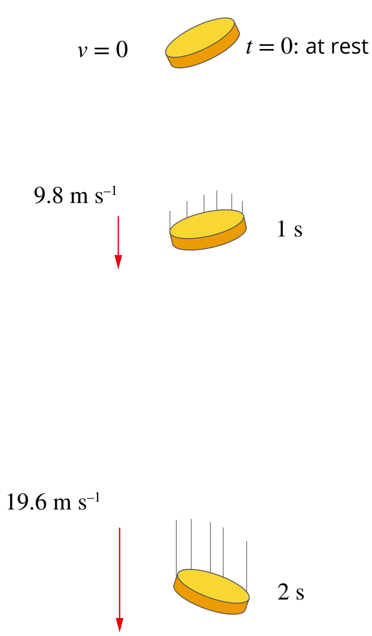


FIGURE 7.5.2 A falling coin.

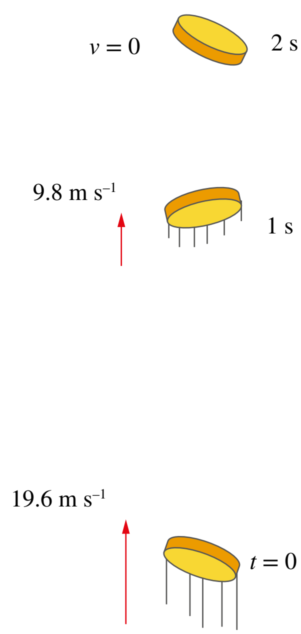


FIGURE 7.5.3 A coin thrown vertically upwards.

PHYSICSFILE

Galileo's experiment on the Moon

In 1971, David Scott went to great lengths to show that Galileo's prediction was correct. As an astronaut on the Apollo 15 Moon mission, he took a hammer and a feather on the voyage. He stepped onto the lunar surface, held the feather and hammer at the same height and dropped them together. As Galileo had predicted 400 years earlier, in the absence of any air resistance the two objects fell side by side as they accelerated towards the Moon's surface.

FIGURE 7.5.4 Astronaut David Scott holding a feather and a hammer on the Moon.

PHYSICS IN ACTION

Theories of motion: Aristotle and Galileo

Aristotle was a Greek philosopher who lived in the fourth century BCE. He was such an influential individual that his ideas on motion were generally accepted for nearly 2000 years. Aristotle did not do experiments as we know them today, but simply *thought* about different bodies in order to arrive at a plausible explanation for their motion.

Aristotle spent a lot of time classifying animals, and adopted a similar approach in his study of motion. His theory gave inanimate objects, such as rocks and rain, similar characteristics to living things. Aristotle organised objects into four terrestrial groups or elements: earth, water, air and fire (see Figure 7.5.5). He said that any object was a mixture of these elements in a certain proportion.

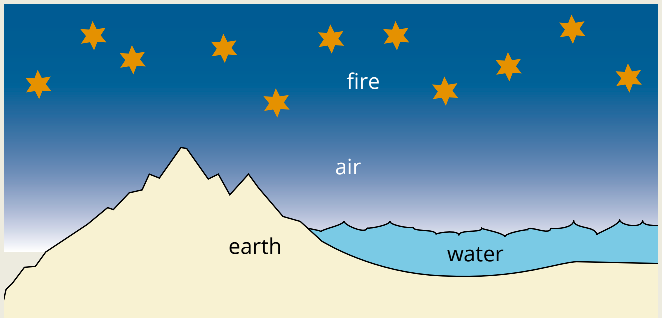


FIGURE 7.5.5 Aristotle's four elements of the universe; earth, water, air and fire.

According to Aristotle, a body would move because of a tendency that could come from inside or outside the body. An internal tendency would cause 'natural' motion and result in a body returning to its proper place. For example, if a rock, which is an earth substance, is held in the air and released, its natural tendency would be to return to Earth. This explains why it falls down. Similarly, fire was thought to head upwards in an attempt to return to its proper place in the universe.

An external push that acts when something is thrown or hit was the cause of 'violent' motion in the Aristotelian model. An external push acted to take a body away from its proper place. For example, when an apple is thrown into the air, a violent motion carries the apple away from the Earth, but then the natural tendency of the apple takes over and it returns to its home.

Aristotle's theory worked quite well and could be used to explain many observed types of motion. However, there were also many examples that it could not successfully explain, such as why some solids floated instead of sinking.

Aristotle explained the behaviour of a falling body by saying that its speed depended on how much earth element it contained. This suggested that a 2 kg cat would fall twice as fast and in half the time as a 1 kg cat dropped from the same height. Many centuries later, Galileo Galilei (pictured in Figure 7.5.6) noticed that, at the start of a hailstorm, small hailstones arrived at the same time as large hailstones. This caused Galileo to doubt Aristotle's theory and so he set about finding an explanation for the motion of freely falling bodies.

FIGURE 7.5.6 Galileo Galilei.

A famous story in science is that of Galileo dropping different weights from the Leaning Tower of Pisa in Italy. This story may or may not be true, but Galileo did perform a very detailed analysis of falling bodies. Galileo used inclined planes because freely falling bodies moved too fast to analyse. He completed extensive and thorough experiments that showed conclusively that Aristotle was incorrect.

By using a water clock to time balls as they rolled down different inclines, he was able to show that the balls were accelerating and that the distance they travelled was proportional to the square of the time, i.e. $d \propto t^2$.

Galileo found that this also held true when he inclined the plane at larger and larger angles, allowing him to conclude that freely falling bodies actually fall with a uniform acceleration.

PHYSICSFILE

Strength of gravity

The acceleration due to gravity, g, on Earth varies slightly from 9.80 m s^{-2} according to the location. The reasons for this will be studied in Physics Unit 3. On the Moon, the strength of gravity, g, is much weaker than on Earth and falling objects accelerate at 1.60 m s^{-2}. Other planets and bodies in the solar system have different values of g depending on their mass and size. The value of g at various locations is provided in Table 7.5.1.

TABLE 7.5.1 Acceleration due to gravity at different locations on Earth, and on other bodies in the solar system.

Location	Acceleration due to gravity (m s^{-2})
Perth	9.794
South Pole	9.832
Equator	9.780
Moon	1.600
Mars	3.600
Jupiter	24.600
Pluto	0.670

In Physics Unit 2 you can assume that acceleration due to gravity is always 9.80 m s^{-2}.

Since the acceleration of a free-falling body is constant, it is appropriate to use the equations that were studied in the previous section, 'Equations for uniform acceleration'. It is necessary to specify whether up or down is positive when doing these problems. You can simply follow the mathematical convention of regarding up as positive, which would mean the acceleration due to gravity would always be −9.80 m s^{-2}.

Worked example 7.5.1

VERTICAL MOTION

A construction worker accidentally knocks a brick from a building so that it falls vertically a distance of 50 m to the ground. Use g = −9.80 m s^{-2} and ignore air resistance when answering these questions.

a How long does the brick take to fall halfway, to 25 m?

Thinking	Working
Write down the known quantities and the quantity that you need to find. Apply the sign convention that up is positive and down is negative.	The brick starts at rest so $u = 0$. $s = -25$ m $u = 0$ m s^{-1} $a = -9.80$ m s^{-2} $t = ?$
Identify the correct equation for uniform acceleration to use.	$s = ut + \frac{1}{2}at^2$
Substitute known values into the equation and solve for t. Think about whether the value seems reasonable.	$-25 = 0 \times t + \frac{1}{2} \times -9.80 \times t^2$ $-25 = -4.90t^2$ $t = \sqrt{\frac{-25}{-4.90}}$ $= 2.3$ s

b How long does the brick take to fall all the way to the ground?

Thinking	Working
Write down the known quantities and the quantity that you need to find. Apply the sign convention that up is positive and down is negative.	$s = -50$ m $u = 0$ m s^{-1} $a = -9.80$ m s^{-1} $t = ?$
Identify the correct equation for uniform acceleration to use.	$s = ut + \frac{1}{2}at^2$
Substitute known values into the equation and solve for t. Think about whether the value seems reasonable. Notice that the brick takes 2.3 s to travel the first 25 m and only 0.9 s to travel the final 25 m. This is because it is accelerating.	$-50 = 0 \times t + \frac{1}{2} \times -9.80 \times t^2$ $-50 = -4.90t^2$ $t = \sqrt{\frac{-50}{-4.90}}$ $= 3.2$ s

c What is the velocity of the brick as it hits the ground?	
Thinking	**Working**
Write down the known quantities and the quantity that you need to find. Apply the sign convention that up is positive and down is negative.	$s = -50\,m$ $u = 0\,m\,s^{-1}$ $v = ?$ $a = -9.80\,m\,s^{-1}$ $t = 3.2\,s$
Identify the correct equation to use. Since you now know four values, any equation involving *v* will work.	$v = u + at$
Substitute the known values into the equation and solve for *v*. Think about whether the value seems reasonable.	$v = 0 + (-9.80) \times 3.2$ $= -31\,m\,s^{-1}$
Use the sign and direction convention to describe the direction of the final velocity.	$v = -31\,m\,s^{-1}$ or $31\,m\,s^{-1}$ downwards

Worked example: Try yourself 7.5.1

VERTICAL MOTION

A construction worker accidentally knocks a hammer from a building so that it falls vertically a distance of 60 m to the ground. Use $g = -9.80\,m\,s^{-2}$ and ignore air resistance when answering these questions.

a How long does the hammer take to fall halfway, to 30 m?

b How long does it take the hammer to fall all the way to the ground?

c What is the velocity of the hammer as it hits the ground?

When an object is thrown vertically up into the air, it will eventually reach a point where it stops momentarily before returning back down. So the velocity of the object decreases as the object rises, becomes zero at the maximum height, and then increases again as the object falls. Throughout this motion, however, the object is still in the same gravitational field, so g remains at $-9.8\,m\,s^{-2}$ throughout the journey. Knowing that the velocity of an object thrown in the air is zero at the top of its flight allows you to calculate the maximum height reached.

Worked example 7.5.2

MAXIMUM HEIGHT PROBLEMS

On winning a tennis match the victorious player, Michael, smashed the ball vertically into the air at 27.5 m s^{-1}. In the following questions, ignore air resistance and use $g = 9.80$ m s^{-2}.

a Determine the maximum height reached by the ball.	
Thinking	**Working**
Write down the known quantities and the quantity that you need to find. At the maximum height the velocity is zero. Apply the sign convention that up is positive and down is negative.	$u = 27.5$ m s^{-1} $v = 0$ $a = -9.80$ m s^{-2} $s = ?$
Identify the correct equation to use.	$v^2 = u^2 + 2as$
Substitute known values into the equation and solve for s.	$0 = (27.5)^2 + 2 \times (-9.80) \times s$ $s = \frac{-756.25}{2 \times (-9.80)} = 38.6$ $\therefore s = +38.6$ m, i.e. the ball reaches a height of 38.6 m.

b Calculate the time that the ball takes to return to its starting position.	
Thinking	**Working**
To work out the time the ball is in the air, first calculate the time it takes to reach its maximum height. Write down the known quantities and the quantity that you need to find.	$u = 27.5$ m s^{-1} $v = 0$ m s^{-1} $a = -9.80$ m s^{-2} $s = 38.6$ $t = ?$
Identify the correct equation to use.	$v = u + at$
Substitute known values into the equation and solve for t.	$0 = 27.5 + (-9.80 \times t)$ $9.80t = 27.5$ $\therefore t = 2.80$ s The ball takes 2.80 s to reach its maximum height. It will therefore take 2.80 s to fall from this height back to its starting point and so it takes 5.60 s to return to its starting position.

Worked example: Try yourself 7.5.2

MAXIMUM HEIGHT PROBLEMS

On winning a cricket match, a fielder throws a cricket ball vertically into the air at 15 m s^{-1}. In the following questions, ignore air resistance and use $g = 9.80$ m s^{-2}.

a Determine the maximum height reached by the ball.

b Calculate the time that the ball takes to return to its starting position.

7.5 Review

SUMMARY

- If air resistance can be ignored, all bodies falling freely near the Earth will move with the same constant acceleration.
- The acceleration due to gravity is represented by g and is equal to $9.80\,m\,s^{-2}$ in the direction towards the centre of the Earth.
- The equations for uniform acceleration can be used to solve vertical motion problems. It is necessary to specify whether up or down is positive.

KEY QUESTIONS

For these questions, ignore the effects of air resistance and assume that the acceleration due to gravity is $9.80\,m\,s^{-2}$ unless instructed otherwise.

1 A ball is thrown into the air. Describe how the velocity of the ball changes when it leaves the hand.

2 Angus inadvertently drops an egg while baking a cake and the egg falls vertically towards the ground. Which one of the following statements correctly describes how the egg falls?
- **A** The egg's acceleration increases.
- **B** The egg's acceleration is constant.
- **C** The egg's velocity is constant.
- **D** The egg's acceleration decreases.

3 Max is an Olympic trampolinist and is practising some routines. Which one or more of the following statements correctly describes Max's motion when he is at the highest point of the bounce? Assume that his motion is vertical.
- **A** He has zero velocity.
- **B** His acceleration is zero.
- **C** His acceleration is upwards.
- **D** His acceleration is downwards.

4 A window cleaner working on the Bell tower accidently drops her mobile phone. The phone falls vertically towards the ground with an acceleration of $9.80\,m\,s^{-2}$.
- **a** Determine the speed of the phone after 3.0 s.
- **b** How fast is the phone moving after it has fallen 30 m?
- **c** What is the average velocity of the phone during a fall of 30 m?

5 A girl tosses a marble straight up into the air at $5\,m\,s^{-1}$ and then catches it at the same height from which it was thrown. Ignore air resistance.
- **a** Is the acceleration of the marble on the way up the same as, less than or greater than its acceleration on the way down? Justify your answer.
- **b** Is the launch speed of the marble the same as, less than or greater than its landing speed? Justify your answer.

6 A super ball is bounced so that it travels straight up into the air, reaching its highest point after 1.5 s.
- **a** What is the initial speed of the ball just as it leaves the ground?
- **b** What is the maximum height reached by the ball?

7 A book is knocked off a bench and falls vertically to the floor. If the book takes 0.40 s to fall to the floor, calculate the following descriptions of its motion.
- **a** What is the book's speed as it lands?
- **b** What is the height from which the book fell?
- **c** How far did the book fall during the first 0.20 s?
- **d** How far did the book fall during the final 0.20 s?

8 While celebrating her birthday, Bindi pops a party popper. The lid travels vertically into the air. Being a keen physics student, Bindi notices that the lid takes 4.0 s to return to its starting position.
- **a** How long does the lid take to reach its maximum height?
- **b** How fast was the lid travelling initially?
- **c** What was the maximum height reached by the lid?
- **d** What was the velocity of the lid as it returned to its starting point?

9 Two physics students conduct the following experiment from a very high bridge. Thao drops a 1.5 kg shot-put from a vertical height of 60.0 m, while at exactly the same time Benjamin throws a 100 g mass with an initial downwards velocity of $10.0\,m\,s^{-1}$ from a point 10.0 m above Thao.
- **a** How long does it take the shot-put to reach the ground?
- **b** How long does it take the 100 g mass to reach the ground?

7.5 Review *continued*

10 At the start of a football match, the umpire bounces the ball so that it travels vertically upwards and reaches a height of 15.0 m.

- **a** How long does the ball take to reach this maximum height?
- **b** One of the ruckmen is able to leap and reach to a height of 4.0 m with his hand. How long after the bounce should this ruckman try to make contact with the ball?

11 A stone is thrown vertically upwards at 8.00 m s^{-1} from the top of a cliff and it lands in the sea 3.00 s later. Calculate

- **a** the maximum height above the cliff reached by the stone
- **b** the time taken by the stone to reach its maximum height
- **c** the height of the cliff above the sea.

Chapter review

KEY TERMS

acceleration
air resistance
centre of mass
dimensional analysis
displacement
distance travelled
free fall
magnitude
position
speed
velocity

For the following questions, ignore air resistance and use $g = 9.80\,m\,s^{-2}$, unless indicated otherwise.

1 A car travels at $95\,km\,h^{-1}$ along a freeway. What is its speed in $m\,s^{-1}$?

2 A cyclist travels at $15\,m\,s^{-1}$ during a sprint finish. What is this speed in $km\,h^{-1}$?

The following information relates to questions 3 and 4.

An athlete in training for a marathon runs 15 km north along a straight road before realising that she has dropped her drink bottle. She turns around and runs back 5 km to find her bottle, then resumes running in the original direction. After running for 3.0 hours, the athlete reaches 20 km from her starting position and stops.

3 Calculate the average speed of the athlete in $km\,h^{-1}$.

4 Calculate her average velocity in:

a $km\,h^{-1}$

b $m\,s^{-1}$.

5 A ping pong ball is falling vertically at $6.0\,m\,s^{-1}$ as it hits the floor. It rebounds at $4.0\,m\,s^{-1}$ up. What is its change in speed during the bounce?

6 A car is moving in a positive direction. It approaches a red light and slows down. Which of the following statements correctly describes its acceleration and velocity as it slows down?

A The car has positive acceleration and negative velocity.

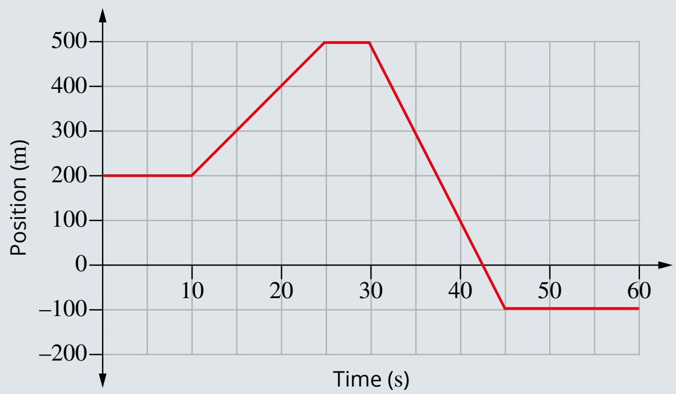
B The car has negative acceleration and positive velocity.

C Both the velocity and acceleration of the car are positive.

D Both the velocity and acceleration of the car are negative.

7 A skier is travelling along a horizontal ski run at a speed of $15\,m\,s^{-1}$. After falling over, the skier takes 2.5 s to come to rest. Calculate the average acceleration of the skier.

8 The graph below shows the position of a motorbike along a straight stretch of road as a function of time. The motorcyclist starts 200 m north of an intersection.



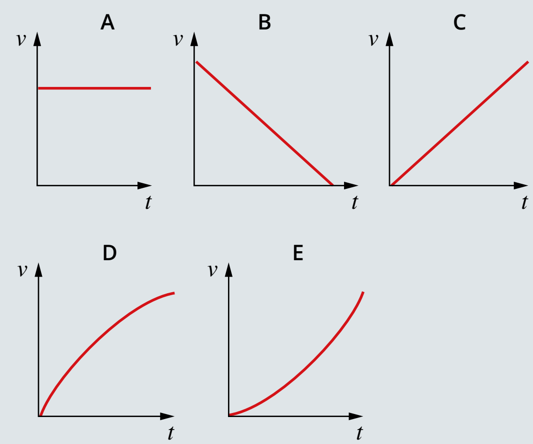
a At what time interval is the motorcyclist travelling in a northerly direction?

b At what time interval is the motorcyclist travelling in a southerly direction?

c At what time intervals is the motorcyclist stationary?

d At what time is the motorcyclist passing back through the intersection?

9 For each of the activities below, indicate which of the following velocity–time graphs best represents the motion involved.

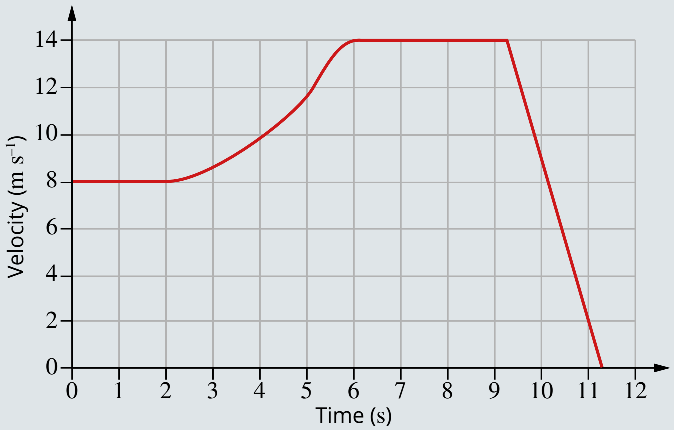


a A car comes to a stop at a red light.

b A swimmer is travelling at a constant speed.

c A motorbike starts from rest with uniform acceleration.

10 This velocity–time graph is for an Olympic road cyclist as he travels, initially north, along a straight section of track.



a Calculate the displacement of the cyclist during his journey.

b Calculate the magnitude, to three significant figures, of the average velocity of the cyclist during this 11.0 s interval.

c Calculate the acceleration of the cyclist at $t = 1$ s.

d Calculate the acceleration of the cyclist at $t = 10$ s.

e Which one or more of the following statements correctly describes the motion of the cyclist?

A He is always travelling north.

B He travels south during the final 2 s.

C He is stationary at $t = 8$ s

D He returns to the starting point after 11 s.

11 A car starts from rest and has a constant acceleration of 3.5 m s^{-2} for 4.5 s. What is its final speed?

12 A jet-ski starts from rest and accelerates uniformly. If it travels 2.0 m in its first second of motion, calculate:

a its acceleration

b its speed at the end of the first second

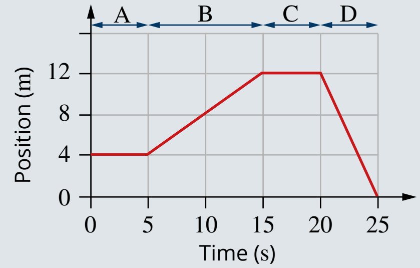
c the distance the jet-ski travels in its second, second of motion.

13 A skater is travelling along a horizontal skate rink at a speed of 10 m s^{-1}. After falling over, the skater takes 10 m to come to rest. Calculate, to two significant figures, the answers to the following questions about the skater's movement.

a What is the average acceleration of the skater?

b How long does it take the skater to come to a stop?

14 The graph shows the position of Candice dancing across a stage.



a What is Candice's starting position?

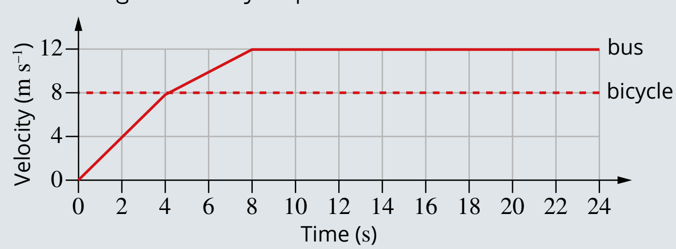
b In which of the sections (A–D) is Candice at rest?

c In which of the sections (A–D) is Candice moving in a positive direction, and what is her velocity?

d In which of the sections (A–D) is Candice moving with a negative velocity and what is the magnitude of this velocity?

e Calculate Candice's average speed during the 25 s motion.

15 The velocity–time graphs for a bus and a bicycle travelling along the same straight stretch of road are shown below. The bus is initially at rest and starts moving as the bicycle passes it.



a What is the magnitude of the initial acceleration of the bus?

b At what time does the bus overtake the bicycle?

c How far has the bicycle travelled before the bus catches it?

d What is the magnitude of the average velocity of the bus during the first 8 s?

16 **a** Draw an acceleration–time graph for the bus discussed in question 11.

b Use your acceleration–time graph to determine the change in velocity of the bus over the first 8 s.

17 A slingshot is used to launch a marble vertically into the air at 39.2 m s^{-1}. Discuss the velocity and acceleration of the marble as it travels to its maximum height. Indicate the time that it takes to reach the top. Consider up as positive.

18 A golfer mis-hits a golf ball straight up into the air. Which one of the following statements best describes the acceleration of the ball while it is in the air?

A The acceleration of the ball decreases as it travels upwards, becoming zero as it reaches its highest point.

B The acceleration is constant as the ball travels upwards, then reverses direction as the ball falls down again.

C The acceleration of the ball is greatest when the ball is at the highest point.

D The acceleration is constant for all the time the ball is in the air.

19 Steph tosses a rock vertically into the air. Which of the options below correctly fills the blanks of the following statement about the rock's motion?

On its way upwards, the rock has ______________________ velocity and ____________________ acceleration. At the highest point, the rock has ___________________ velocity and ___________________ acceleration. On its way downwards, the rock has __________________ velocity and ____________________ acceleration.

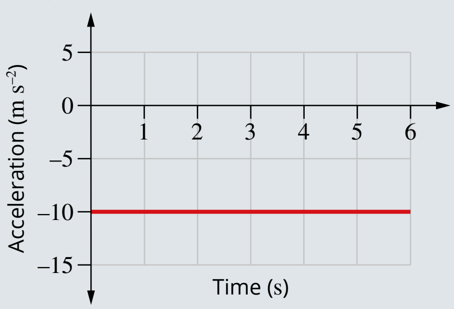
A upwards; upwards; zero; downwards; downwards; downwards

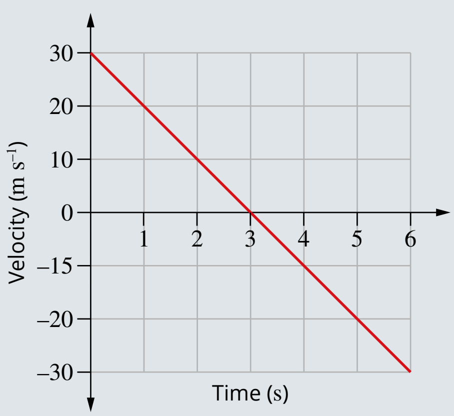
B upwards; downwards; zero; downwards; downwards; downwards

C upwards; upwards; zero; zero; downwards; downwards

D upwards; downwards; zero; zero; downwards; downwards

20 After winning a tennis match, Claire hit a tennis ball vertically into the air at 30 m s^{-1}. The *v–t* and *a–t* graphs for the tennis ball are shown below. Use the graphs or the equations for uniform acceleration to answer the following questions. Use $g = 10\,\text{m s}^{-2}$ for these questions. Assume the motion in question is symmetrical, starting and ending at the same point.





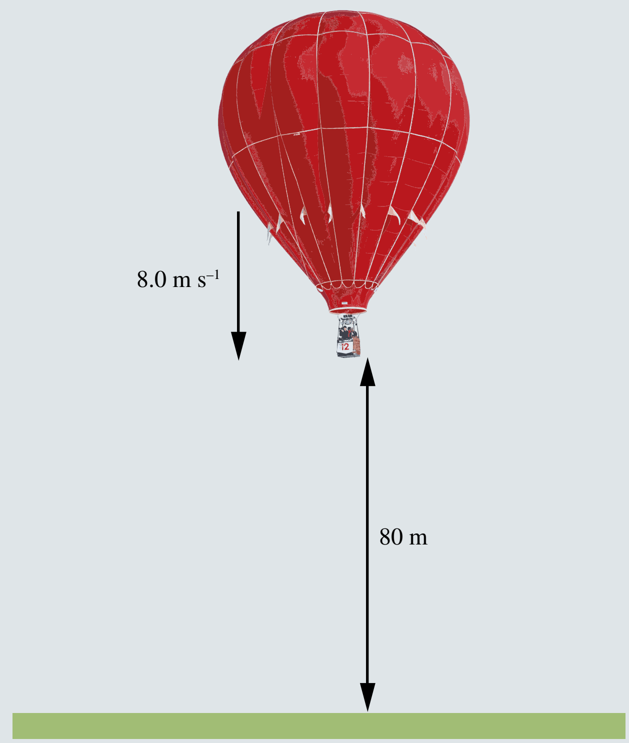
a What is the maximum height reached by the ball?

b What is the time that the ball takes to return to its starting position?

c What is the velocity of the ball 5.0 s after Claire hits it?

d What is the acceleration of the ball at its maximum height?

21 A hot-air balloon is 80 m above the ground and travelling vertically downwards at 8.0 m s^{-1} when one of the passengers, Olivia, accidentally drops a coin over the side.



a How long does the balloon take to reach the ground?

b What is the speed of the coin as it reaches the ground?

c How long after the coin reaches the ground does the balloon touch down?

The following information relates to questions 22 and 23.

During a game of minigolf, Renee putts a ball so that it hits an obstacle and travels straight up into the air, reaching its highest point after 1.5 s.

22 What was the initial velocity of the ball as it launched into the air?

23 Calculate the maximum height reached by the ball.

CHAPTER

08 Momentum and forces

In the seventeenth century, Sir Isaac Newton published three laws that explain why objects in our universe move as they do. These laws became the foundation of a branch of physics called mechanics: the science of how and why objects move. They have become commonly known as Newton's three laws of motion.

Using Newton's laws, this chapter will describe the relationship between the forces acting on an object and its motion. You will also learn about the relationship between force, period of time and change in momentum (impulse).

Science as a Human Endeavour

Safety for motorists and other road users has been substantially increased through application of Newton's laws and conservation of momentum by the development and use of devices, including:

- helmets
- seatbelts
- crumple zones
- airbags
- safety barriers

Science Understanding

- Momentum is a property of moving objects; it is conserved in a closed system and may be transferred from one object to another when a force acts over a time interval. This includes applying the relationships:

  $p = mv$, $\Sigma mv_{\text{before}} = \Sigma mv_{\text{after}}$, $mv - mu = \Delta p = F\Delta t$

- Newton's three laws of motion describe the relationship between the force or forces acting on an object, modelled as a point mass, and the motion of the object due to the application of the force or forces.
- free body diagrams show the forces and net force acting on objects, from descriptions of real-life situations involving forces acting in one or two dimensions This includes applying the relationships:

  resultant $F = ma$, $F_{\text{weight}} = mg$

8.1 Momentum and conservation of momentum

It is possible to understand some physical concepts intuitively without knowing the physics terms or words that describe them. For example, you may know that once a heavy object gets moving it is difficult to stop it, whereas a lighter object moving at the same speed is easier to stop. In the next sections of this chapter you will see how Newton's laws of motion can be used to explain these observations. In this section you will explore how these observations can be related to the concept called momentum.

MOMENTUM

The **momentum** of an object relates to both its mass and its velocity. The footballers colliding in Figure 8.1.1 have momentum due to their masses and velocities. The faster they run, the more momentum they will have. A slower-moving player has less momentum than one who is moving faster. Similarly, a person with greater mass will have more momentum than a smaller, lighter person travelling at the same speed. The more momentum an object has (due to its mass or velocity) the more momentum it has to lose before it stops.

FIGURE 8.1.1 Momentum is related to mass and velocity. The greater the mass or velocity, the harder it is to stop or start moving.

The momentum, p, of an object is the product of its mass, m, and its velocity, v.

$p = mv$

where p is momentum (kg m s^{-1})

m is the mass of the object (kg)

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

The greater an object's mass or velocity, the larger that object's momentum will be. As velocity is a vector quantity, momentum is also a vector and so it must have magnitude, direction and units. The direction of a momentum vector will always be the same as the direction of the velocity vector. For calculations of change in momentum in a single dimension, use the sign conventions of positive and negative.

As you will see in Section 8.4, force is equal to the rate of change of momentum. This means that any changes in momentum are caused by the action of an external net force.

Worked example 8.1.1

MOMENTUM

Calculate the momentum of a 60.0 kg student walking at 3.5 m s^{-1} east.	
Thinking	**Working**
Ensure that the variables are in their standard units.	$m = 60.0$ kg $v = 3.50$ m s^{-1} east
Apply the equation for momentum.	$p = mv$ $= 60.0 \times 3.50$ $= 210$ kg m s^{-1}
Ensure that the final answer is in the same direction as the velocity.	$p = 210$ kg m s^{-1} east

Worked example: Try yourself 8.1.1

MOMENTUM

Calculate the momentum of a 1230 kg car driving at 16.7 m s^{-1} north.

CONSERVATION OF MOMENTUM

The most significant feature of momentum is that it is **conserved** in any interaction or collision between objects. This means that the total (sum of) momentum in any system before a collision will be equal to the total (sum of) momentum in the system after the collision. This is known as the law of conservation of momentum and can be represented by the following relationship:

$$\Sigma p_{\text{before}} = \Sigma p_{\text{after}}$$

where Σp is the sum of the momentum of objects in a system.

Or:

$$\Sigma mv_{\text{before}} = \Sigma mv_{\text{after}}$$

To find the total momentum of objects in a system (either before or after a collision) simply find the momentum of each object, using their masses and velocities, and then add them together (making sure to perform the appropriate vector addition).

For collisions in one dimension, apply the sign convention of positive and negative directions to the velocities and then use algebra to determine the answer to the problem. For collisions in two dimensions, resolve vectors describing motion into perpendicular components and then consider the conservation of momentum in each single dimension.

Momentum in one-dimensional collisions

If two objects are colliding in one dimension, then the following equation applies:

$$\Sigma mv_{\text{before}} = \Sigma mv_{\text{after}}$$

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_1$$

where m_1 is the mass of object 1 (kg)

u_1 is the initial velocity of object 1 (m s^{-1})

v_1 is the final velocity of object 1 (m s^{-1})

m_2 is the mass of object 2 (kg)

u_2 is the initial velocity of object 2 (m s^{-1})

v_2 is the final velocity of object 2 (m s^{-1}).

PHYSICSFILE

The discovery of the neutron

The law of conservation of momentum was used to interpret the data from investigations that led to the discovery of the neutron. Because the neutron has no charge, it could not be investigated through the interactions of charged particles that had led to the discovery of the proton and electron. In 1932, James Chadwick investigated collisions between alpha particles and the element beryllium. However, the conservation of momentum calculations didn't add up. Chadwick knew that the law of conservation of momentum was true, so he reasoned that there was an unknown particle involved that had a mass close to the proton's mass, but without electric charge. Subsequent investigations confirmed his experiments and led to the naming of this particle as the neutron.

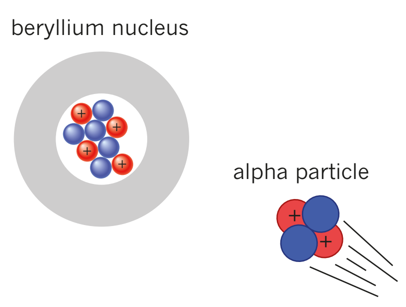


FIGURE 8.1.2 Investigating collisions between alpha particles and the beryllium nucleus led to the discovery of the neutron.

PHYSICS IN ACTION

Elastic and inelastic collisions

Collisions can either be elastic or inelastic. In elastic collisions no kinetic energy is lost, while inelastic collisions will result in some kinetic energy being converted to another form, such as thermal energy. The loss in kinetic energy during inelastic collisions will result in a lower combination of final velocities. If the collision was 100% inelastic, then all of the kinetic energy would be converted to thermal energy and both objects would be stationary immediately after the collision. If the collision was 100% elastic, then the final velocities would be maximum.

It is impossible to predict the final velocities of two objects colliding unless the collision is perfectly elastic, as it is not possible to predict the kinetic energy lost in any collision. Most collisions between two objects in real life are inelastic. However, some large-scale interactions—such as a gravitational slingshot between a rocket and a planet—are perfectly elastic (Figure 8.1.3). You will explore this concept more in Section 9.3.

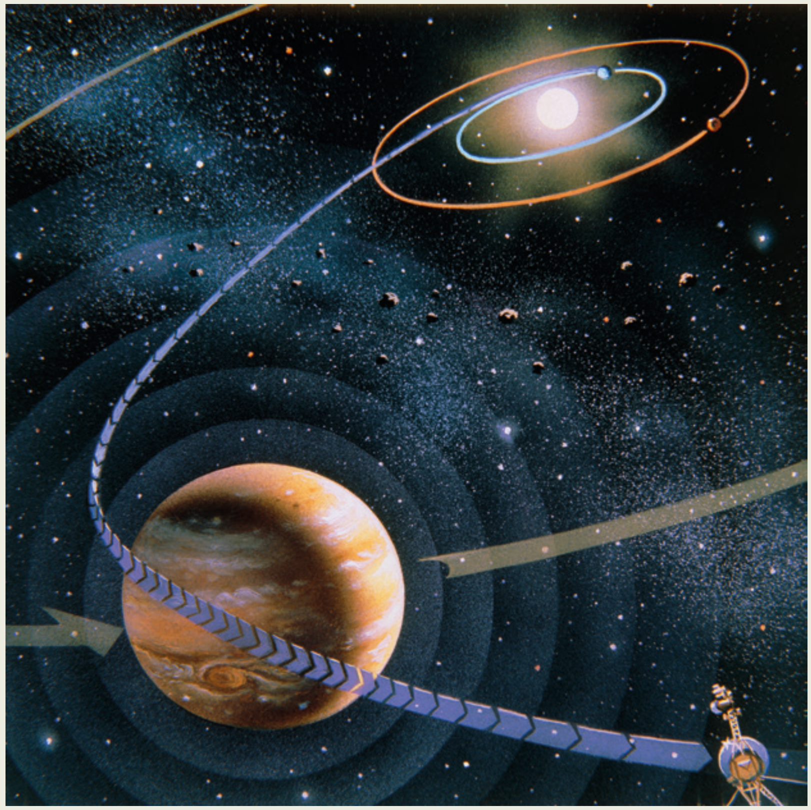


FIGURE 8.1.3 Artist's impression of the gravity assist flyby, also known as gravitational slingshot, of the spacecraft Voyager 2 around Jupiter.

Worked example 8.1.2

CONSERVATION OF MOMENTUM

A 2.50 kg mass is moving at 4.50 m s^{-1} west towards a 1.50 kg mass moving at 3.00 m s^{-1} east. Calculate the final velocity of the 2.50 kg mass if the 1.50 kg mass rebounds at 5.00 m s^{-1} west.

Thinking	Working
Identify the variables using subscripts. Ensure that the variables are in their standard units.	$m_1 = 2.50$ kg $u_1 = 4.50$ m s^{-1} west $v_1 = ?$ $m_2 = 1.50$ kg $u_2 = 3.00$ m s^{-1} east $v_2 = 5.00$ m s^{-1} west
Apply the sign convention to the variables. In this case east is positive and west is negative.	$m_1 = 2.50$ kg $u_1 = -4.50$ m s^{-1} $v_1 = ?$ $m_2 = 1.50$ kg $u_2 = +3.00$ m s^{-1} $v_2 = -5.00$ m s^{-1}

Apply the equation for conservation of momentum.	$\Sigma p_{before} = \Sigma p_{after}$ $m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2$ $(2.50 \times -4.50) + (1.50 \times 3.00)$ $= 2.50v_1 + (1.50 \times -5.00)$ $2.50v_1 = -11.25 + 4.50 - (-7.50)$ $v_1 = \frac{0.75}{2.50}$ $= 0.30\,\text{m s}^{-1}$
Apply the sign convention to describe the direction of the final velocity.	$v_1 = 0.30\,\text{m s}^{-1}$ east

Worked example: Try yourself 8.1.2

CONSERVATION OF MOMENTUM

A 1200 kg wrecking ball is moving at $2.50\,\text{m s}^{-1}$ north towards a 1500 kg wrecking ball moving at $4.00\,\text{m s}^{-1}$ south. Calculate the final velocity of the 1500 kg ball if the 1200 kg ball rebounds at $3.50\,\text{m s}^{-1}$ south.

PHYSICS IN ACTION

Conservation of momentum in sports

The law of conservation of momentum has applications in many spheres of human endeavour. In sport, however, the law is particularly evident. Consider the following sports:

- curling
- lawn bowls
- tenpin bowling
- bocce
- pool
- snooker
- marbles.

All of these sports require the athlete to cause one object to collide with the other. The best players are able to control the initial velocity of the striking object such that the magnitude and direction of momentum for both the striking object and the target object after the collision result in a good score or a positional advantage for themselves or their team.

FIGURE 8.1.4 Curling is a sport that uses the conservation of momentum.

FIGURE 8.1.5 Controlling the direction and momentum of this bowling ball will enable to player to pick up a spare.

Momentum when masses combine

It is important to note that in the situations described in Worked example 8.1.2 (p. xxx), the two objects remain separate from each other. However, it is possible for two objects to combine (stick together) when they collide. If two objects combine when they collide, then the equation is modified to:

$$\Sigma p_{before} = \Sigma p_{after}$$
$$\Sigma mv_{before} = \Sigma mv_{after}$$
$$m_1u_1 + m_2u_2 = m_3v_3$$

where m_1 is the mass of object 1 (kg)

u_1 is the initial velocity of object 1 (m s^{-1})

m_2 is the mass of object 2 (kg)

u_2 is the initial velocity of object 2 (m s^{-1})

m_3 is the combined mass of m_1 and m_2 (kg)

v_3 is the final velocity of combined mass of m_1 and m_2 (m s^{-1}).

Worked example 8.1.3

CONSERVATION OF MOMENTUM WHEN MASSES COMBINE

A 5.00 kg lump of clay is moving at 2.00 m s^{-1} west towards a 7.50 kg mass of clay moving at 3.00 m s^{-1} east. They collide to form a single, combined mass of clay. Calculate the final velocity of the combined mass of clay.

Thinking	Working
Identify the variables using subscripts and ensure that the variables are in their standard units. Add m_1 and m_2 to get m_3.	m_1 = 5.00 kg u_1 = 2.00 m s^{-1} west m_2 = 7.50 kg u_2 = 3.00 m s^{-1} east m_3 = 12.50 kg v_3 = ?
Apply the sign convention to the variables. In this case east is positive and west is negative.	m_1 = 5.00 kg u_1 = –2.00 m s^{-1} m_2 = 7.50 kg u_2 = +3.00 m s^{-1} m_3 = 12.50 kg v_3 = ?
Apply the equation for conservation of momentum.	$\Sigma p_{before} = \Sigma p_{after}$ $m_1u_1 + m_2u_2 = m_3v_3$ $(5.00 \times -2.00) + (7.50 \times 3.00) = 12.50v_3$ $v_3 = \frac{-10.0 + 22.50}{12.50}$ $= 1.00$ m s^{-1}
Apply the sign convention to describe the direction of the final velocity.	v_3 = 1.00 m s^{-1} east

Worked example: Try yourself 8.1.3

CONSERVATION OF MOMENTUM WHEN MASSES COMBINE

An 80.0 kg rugby player is moving at 1.50 m s^{-1} north when he tackles an opponent with a mass of 50.0 kg who is moving at 5.00 m s^{-1} south. Calculate the final velocity of the two players.

Momentum in explosive collisions

It is also possible for one object to break apart into two objects in what is known as an 'explosive collision'. If an object breaks apart when an explosive collision occurs, then the equation is modified to:

$$\Sigma p_{before} = \Sigma p_{after}$$
$$\Sigma mv_{before} = \Sigma mv_{after}$$
$$m_1u_1 = m_2v_2 + m_3v_3$$

where m_1 is the mass of object 1 (2 and 3 combined; kg)
u_1 is the initial velocity of object 1 ($m\,s^{-1}$)
m_2 is the mass of object 2 (kg)
v_2 is the final velocity of object 2 ($m\,s^{-1}$)
m_3 is the mass of object 3 (kg)
v_3 is the final velocity of object 3 ($m\,s^{-1}$).

Worked example 8.1.4

CONSERVATION OF MOMENTUM FOR EXPLOSIVE COLLISIONS

A 90.0 kg athlete holds a 1000 g javelin. She approaches the line at 7.75 $m\,s^{-1}$ west and releases the javelin down the field. After throwing it, she continues with a velocity of 7.25 $m\,s^{-1}$ west. Calculate the velocity of the javelin just after she releases it.

Thinking	Working
Identify the variables using subscripts and ensure that the variables are in their standard units. Note that m_1 is the sum of the bodies, i.e. the athlete and the javelin.	$m_1 = 91$ kg $u_1 = 7.75\ m\,s^{-1}$ west $m_2 = 90$ kg $v_2 = 7.25\ m\,s^{-1}$ west $m_3 = 1.00$ kg $v_3 = ?$
Apply the sign convention to the variables.	$m_1 = 91$ kg $u_1 = -7.75\ m\,s^{-1}$ $m_2 = 90$ kg $v_2 = -7.25\ m\,s^{-1}$ $m_3 = 1.00$ kg $v_3 = ?$
Apply the equation for conservation of momentum for explosive collisions.	$\Sigma p_{before} = \Sigma p_{after}$ $m_1u_1 = m_2u_2 + m_3v_3$ $91.0 \times -7.75 = (90.0 \times -7.25) + 1.00v_3$ $v_3 = \frac{-705.25 - (-652.5)}{1.00}$ $v_3 = \frac{-52.75}{1.00} = -52.8\ m\,s^{-1}$
Apply the sign convention to describe the direction of the final velocity.	$v_3 = 52.8\ m\,s^{-1}$ west

Worked example: Try yourself 8.1.4

CONSERVATION OF MOMENTUM FOR EXPLOSIVE COLLISIONS

A 2000 kg cannon fires a 10 kg cannonball. The cannon and the cannonball are initially stationary. After firing, the cannon recoils with a velocity of 8.15 $m\,s^{-1}$ north. Calculate the velocity of the cannonball just after it is fired.

PHYSICSFILE

Conservation of momentum in engines

If you release an inflated rubber balloon with its neck open, it will fly off around the room. In the diagram below, the momentum of the air to the left results in the movement of the balloon to the right. Momentum is conserved.

This is the principle upon which rockets and jet engines are based. Both rockets and jet engines employ a high-velocity stream of hot gases that are vented after the combustion of a fuel-air mixture. The hot exhaust gases have a very large momentum as a result of the high velocities involved, and can accelerate rockets and jets to high velocities as they acquire an equal momentum in the opposite direction. Rockets destined for space carry their own oxygen supply, while jet engines use the surrounding air supply.

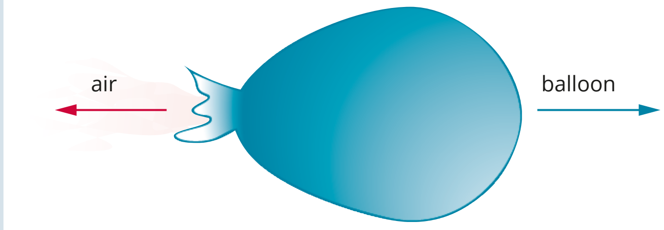


FIGURE 8.1.6 Conservation of momentum demonstrated by an open balloon.

8.1 Review

SUMMARY

- Momentum is the product of an object's mass and velocity.
- Momentum is a vector quantity and is calculated using the equation $p = mv$.
- Force is equal to the rate of change of momentum.
- The law of conservation of momentum can be applied to situations in which:
  - two objects collide and remain separate:
    $\Sigma p_{\text{before}} = \Sigma p_{\text{after}}$
    $\Sigma mv_{\text{before}} = \Sigma mv_{\text{after}}$
    $m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2$
  - two objects collide and combine together:
    $\Sigma p_{\text{before}} = \Sigma p_{\text{after}}$
    $\Sigma mv_{\text{before}} = \Sigma mv_{\text{after}}$
    $m_1u_1 + m_2u_2 = m_3v_3$
  - one object breaks apart into two objects in an explosive collision:
    $\Sigma p_{\text{before}} = \Sigma p_{\text{after}}$
    $\Sigma mv_{\text{before}} = \Sigma mv_{\text{after}}$
    $m_1u_1 = m_2v_2 + m_3v_3$

KEY QUESTIONS

1 Calculate the momentum of a 3.50 kg fish swimming at 2.50 m s^{-1} south.

2 Calculate the momentum of a 433 kg boat travelling at 22.2 m s^{-1} west.

3 Calculate the momentum of a 65.0 g tennis ball being served at 61.0 m s^{-1} south.

4 Which object has the greater momentum: a medicine ball of mass 4.5 kg travelling at 3.5 m s^{-1} or one of mass 2.5 kg travelling at 6.8 m s^{-1}?

5 A rower of mass 70.0 kg steps out of a stationary boat with a velocity of 2.50 m s^{-1} forwards onto the nearby riverbank. The boat has a mass of 400 kg and was initially at rest. With what velocity does the boat begin to move as the rower steps out? Give your answer to three significant figures.

6 A golf ball of mass 70.0 g is stationary on the ground when it is hit by a 545 g golf club travelling at 80.0 m s^{-1}. If the ball leaves the club at a speed of 75.0 m s^{-1}, with what speed does the club move just after hitting the ball? Give your answer to three significant figures.

7 A railway wagon of mass 2.50 tonnes moving along a horizontal track at 2.00 m s^{-1} runs into a stationary engine and is coupled to it. After the collision, the engine and wagon move off together at a slow 0.300 m s^{-1}. What is the mass of the engine alone? Give your answer to three significant figures.

8 A space shuttle of mass 10 000 kg (Including fuel), initially at rest, burns 5.0 kg of fuel and oxygen in its rockets to produce exhaust gases ejected at a velocity of 6000 m s^{-1}. Calculate the velocity that this exchange will give to the space shuttle.

9 A small research rocket of mass 250 kg (including fuel) is launched vertically as part of a weather study. It sends out 50 kg of burnt fuel and exhaust gases with a velocity of 180 m s^{-1} in a 2 s initial acceleration period.

a Calculate the average force that accelerates the exhaust gas.

b Determine the force of the exhaust gas on the rocket.

c Calculate the net upwards acceleration acting on the rocket. (Use $g = 9.80$ m s^{-2} if required.)

d Calculate the velocity of the rocket after this initial acceleration.

8.2 Change in momentum and impulse

The previous section described the momentum of an object in terms of its velocity and its mass. For each of the different collisions described in that section, the momentum of the system was conserved. That is, when all of the objects involved in the collision were considered, the total momentum before and after the collision was the same. But for each separate object, considered in isolation, momentum may not have been conserved. In the examples explored, an object experienced a change in its velocity due to the collision.

When an object changes its velocity, its momentum will also change. An increase in velocity means an increase in momentum, while a decrease in velocity means a decrease in momentum. Change in momentum, Δp, is also called **impulse**, I.

You can also consider this change in momentum as a transfer of momentum. For this reason, momentum transfer is often referred to as impulse, I.

CHANGE IN MOMENTUM IN ONE DIMENSION

It is easy to change the velocity of an object. You can either run faster or run slower; you can press a little harder on the pedals of a bike or press a little softer. You can also bounce an object off a surface. For example, the basketball in Figure 8.2.1 experiences a change in velocity when it changes direction during the bounce. The cause of these changes in motion will be discussed in Section 8.6. First, consider impulse or change in momentum in one dimension.

The term 'impulse' means change in momentum. So the impulse or change in momentum of an object moving in one dimension is calculated using the equation:

$$\begin{aligned} I &= \Delta p \\ &= p_{\text{final}} - p_{\text{initial}} \\ &= mv - mu \end{aligned}$$

where I is the impulse (kg m s^{-1})

Δp is the change in momentum (kg m s^{-1})

m is the mass (kg)

v is the final velocity (m s^{-1})

u is the initial velocity (m s^{-1}).

As momentum is a vector quantity, the impulse or change in momentum is also a vector, so it is expressed in magnitude, units and direction.

FIGURE 8.2.1 A bouncing basketball undergoes a change in momentum when it changes direction as it bounces.

Worked example 8.2.1

IMPULSE OR CHANGE IN MOMENTUM

A student rides a bike to school and approaches the bike rack at 8.20 m s^{-1} east. Calculate the impulse of the student during the time it takes to stop if the student and the bike have a combined mass of 80 kg and the student stops at the rack.

Thinking	Working
Ensure that the variables are in their standard units.	$m = 80$ kg $u = 8.20$ m s^{-1} east $v = 0$ m s^{-1}
Apply the sign convention to the velocity vector.	$m = 80$ kg $u = +8.20$ m s^{-1} $v = 0$ m s^{-1}

Apply the equation for impulse or change in momentum.	$I = mv - mu$ $= (80 \times 0) - (80 \times 8.20)$ $= 0 - 656$ $= -656\,\text{kg m s}^{-1}$
Apply the sign convention to describe the direction of the impulse.	$I = 656\,\text{kg m s}^{-1}$ west

Worked example: Try yourself 8.2.1

IMPULSE OR CHANGE IN MOMENTUM

A student hurries to class after lunch, moving at $4.55\,\text{m s}^{-1}$ north. Suddenly the student remembers that she has forgotten her laptop and goes back to her locker at $6.15\,\text{m s}^{-1}$ south. If her mass is 75.0 kg, calculate the impulse of the student during the time it takes to turn around.

CHANGE IN MOMENTUM IN TWO DIMENSIONS

FIGURE 8.2.2 Changing momentum in two dimensions by changing direction.

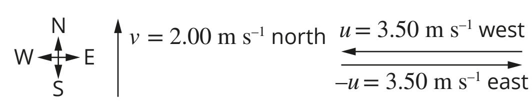
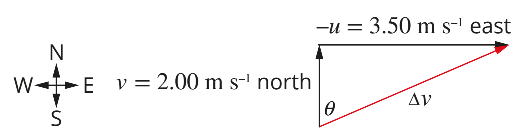
The velocity of an object can be changed not only by changing the magnitude of its velocity, but also by changing the direction of its motion. The velocity of the boat in Figure 8.2.2, for example, changes because the boat changes direction. As you saw in Chapter 6 'Scalars and vectors', a change in velocity in two dimensions can be calculated using geometry. The equation for impulse can be manipulated slightly to illustrate where the change in velocity is applied:

$$I = mv - mu$$
$$= m(v - u)$$

Worked example 8.2.2

IMPULSE OR CHANGE IN MOMENTUM IN TWO DIMENSIONS

A 65.0 kg mass is moving at $3.50\,\text{m s}^{-1}$ west and then changes to $2.00\,\text{m s}^{-1}$ north. Calculate the change in momentum of the mass over the period of the change.

Thinking	Working
Identify the formula for calculating a change in velocity, Δv.	Δv = final velocity – initial velocity $\Delta v = v - u$ $\Delta v = v + (-u)$
Draw the final velocity vector, v, and the initial velocity vector, u, separately. Then draw the initial velocity in the opposite direction, which represents the negative of the initial velocity, $-u$.	
Construct a vector diagram drawing v first and then from its head draw the opposite of u. The change of velocity vector is drawn from the tail of the final velocity to the head of the opposite of the initial velocity.	
As the two vectors to be added are at 90° to each other, apply Pythagoras' theorem to calculate the magnitude of the change in velocity.	$\Delta v^2 = 2.00^2 + 3.50^2$ $= 4.00 + 12.25$ $\Delta v = \sqrt{16.25}$ $= 4.03\,\text{m s}^{-1}$

Calculate the angle from the north vector to the change in velocity vector.	$\tan\theta = \frac{3.50}{2.00}$ $\theta = \tan^{-1} 1.75$ $= 60.3°$
State the magnitude and direction of the change in velocity.	$\Delta v = 4.03\,\text{m s}^{-1}$ N 60.3°E
Identify the variables using subscripts and ensure that the variables are in their standard units.	$m = 65.0\,\text{kg}$ $\Delta v = 4.03\,\text{m s}^{-1}$ N 60.3°E
Apply the equation for impulse or change in momentum.	$\Delta p = mv - mu$ $= m(\text{v} - \text{u})$ $= m\Delta v$ $= 65.0 \times 4.03$ $= 262\,\text{kg m s}^{-1}$
Apply the direction convention to describe the direction of the change in momentum.	$\Delta p = 262\,\text{kg m s}^{-1}$ N 60.3°E

Worked example: Try yourself 8.2.2

IMPULSE OR CHANGE IN MOMENTUM IN TWO DIMENSIONS

A 65.0 g pool ball is moving at $0.250\,\text{m s}^{-1}$ south towards a cushion and bounces off at $0.200\,\text{m s}^{-1}$ east. Calculate the impulse on the ball during the change in velocity.

8.2 Review

SUMMARY

- Change or transfer in momentum, Δp, is also known as impulse, I. It is a vector quantity.
- A change or transfer in momentum occurs when an object changes its velocity.
- The equation for impulse is: $I = \Delta p = mv - mu$
- Change in momentum in two directions can be calculated using geometry.

KEY QUESTIONS

1. Calculate the impulse of a 9.50 kg dog that changes its velocity from $2.50\,\text{m s}^{-1}$ north to $6.25\,\text{m s}^{-1}$ south.
2. Calculate the impulse of a 6050 kg truck as it changes from moving at $22.2\,\text{m s}^{-1}$ west to $16.7\,\text{m s}^{-1}$ east.
3. The velocity of an 8.00 kg mass changes from $3.00\,\text{m s}^{-1}$ east to $8.00\,\text{m s}^{-1}$ east. Calculate the change in momentum.
4. Calculate the change in momentum of a 250 g apple as it changes from rest to moving downwards at $9.80\,\text{m s}^{-1}$ after falling off a tree.
5. The momentum of a ball of mass 0.125 kg changes by $0.075\,\text{kg m s}^{-1}$ south. If its original velocity was $3.00\,\text{m s}^{-1}$ north, what is the final velocity?
6. A 45.0 kg mass moving at $45.0\,\text{m s}^{-1}$ west changes direction so that it moves at $45.0\,\text{m s}^{-1}$ north. Calculate the change in momentum of the mass over the period of the change.
7. A marathon runner with a mass of 70.0 kg is running with a velocity of $4.00\,\text{m s}^{-1}$ north, and then turns a corner to start running $3.60\,\text{m s}^{-1}$ west. Calculate the marathon runner's change in momentum.

8.3 Newton's first law

The previous sections developed the concepts and ideas needed to describe the properties and motion of a moving body. In the following sections, however, rather than simply describing the motion, you will investigate the forces that cause the motion to occur. In addition, you will explore Newton's three laws of motion, which describe and predict the relationship between the forces acting on an object (modeled as a point mass) and its subsequent motion. In this section you will learn about Newton's first law.

FORCE

In simple terms, a **force** can be thought of as a push or a pull, but forces exist in a wide variety of situations in your life and are fundamental to the nature of matter and the structure of the universe. Consider each of the images in Figure 8.3.1. For each situation a force—a push or pull—is acting.

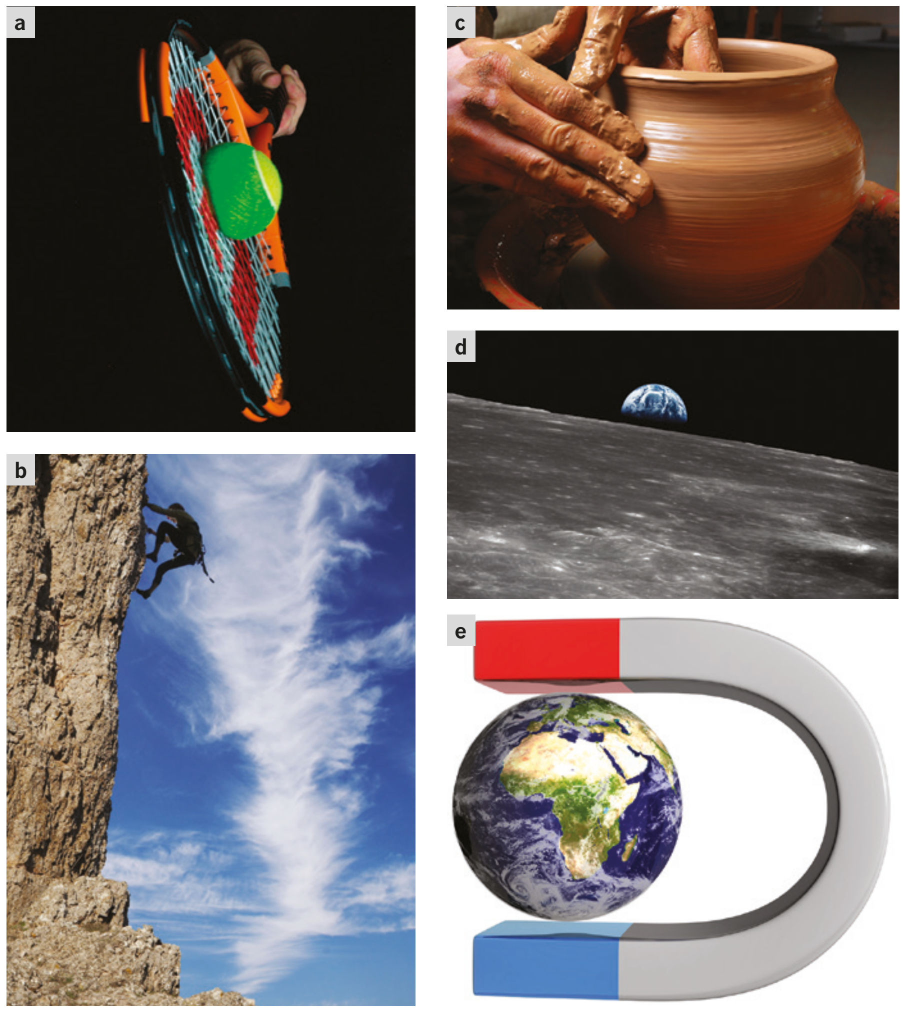


FIGURE 8.3.1 (a) At the moment of impact, both the tennis ball and the racquet strings are distorted by the forces acting at this instant. (b) The rock climber is relying on the frictional force between his hands and feet and the rock face. (c) A continual force causes the clay to deform into the required shape. (d) The gravitational force between the Earth and the Moon is responsible for two high tides each day. (e) The globe is suspended in mid-air because of the magnetic forces of repulsion and attraction.

Some are forces depicted in Figure 8.3.1 are applied directly to an object and some act on a body without touching it. Forces that act directly on a body are called **contact forces**, because the body will only experience the force while contact is maintained. Forces that act on a body at a distance are **non-contact forces**.

Contact forces are the easiest to understand and include the simple pushes and pulls that are experienced daily in people's lives. Examples of these include the forces between colliding billiard balls and the forces that act between you and your chair as you sit reading this book. Friction and drag forces are also contact forces.

Non-contact forces occur when the object causing the push or pull is physically separated from the object that experiences the force. These forces are said to 'act at a distance'. Gravitation, magnetic and electric forces are examples of non-contact forces.

A force measure of the push or pull on an object, and is measured in newtons (N). It is a vector and so it requires a magnitude and a direction to describe it fully.

The amount of force acting can be measured using the SI unit called the newton, which is given the symbol N. The unit, which will be defined later in the chapter, honours Sir Isaac Newton (1642–1727), who is considered to be one of the most significant physicists to have lived and whose first law is the subject of this section. A force of one newton, 1 N, is approximately the force you have to exert when holding a 100 g mass against the downwards pull of gravity. In everyday life this is about the same as holding a small apple.

If more than one force acts on a body at the same time, the body behaves as if only one force—the vector sum of all the forces—is acting. The vector sum of the forces is called the resultant or **net force**, F_{net}. (Note: vectors were covered in detail in Chapter 6.)

The net force acting on a body experiencing a number of forces acting simultaneously is given by the vector sum of all the individual forces:

$$F_{net} = F_1 + F_2 + \ldots + F_n$$

NEWTON'S FIRST LAW

Inertia and Newton's first law are closely related; in fact, some people call Newton's first law the law of inertia. Inertia is the tendency of an object to maintain its velocity. This tendency is related to the mass of an object, so that the greater the mass, the harder it is to get it moving or to stop it from moving.

Newton's first law is a law that is often misunderstood due to a common misconception. People mistakenly think that an object that is moving at constant velocity must have a force causing it to move. This section will address this misconception and will enable you to understand how Newton's first law applies to all situations in which an object moves.

Newton's first law can be stated as:

An object will maintain a constant velocity unless an unbalanced, external force acts on it.

This statement needs to be analysed in more detail by first examining some of the key terms used. The term 'maintain a constant velocity' implies that, if the object is moving, then it will continue to move with a velocity that has the same magnitude and direction. For example, if a car is moving at $12.0\,m\,s^{-1}$ south, then some time later it will still be moving at $12.0\,m\,s^{-1}$ south (Figure 8.3.2). It should also be noted that zero velocity can also be constant, so if the car is moving at $0\,m\,s^{-1}$, then some time later it will still be moving at $0\,m\,s^{-1}$.

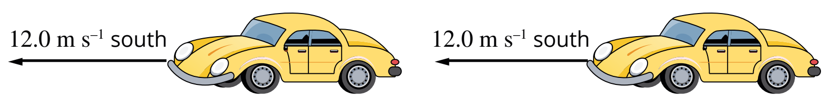


FIGURE 8.3.2 A car maintaining a constant velocity.

The term 'unless' is particularly important in Newton's statement. In effect, it tells you what must *not* happen for the motion to be constant instead of telling you what must happen. Rephrasing Newton's first law to remove the term 'unless' means that it could say:

An object will not maintain its velocity if an unbalanced, external force is applied.

In this context, *not* continuing with its velocity implies that the object is changing its velocity. A change in velocity means that the object is accelerating.

The use of the term 'unbalanced' in relation to the acting force implies that there must be a net force acting on the object. If the forces are balanced, then the object's velocity will remain constant. If the forces are unbalanced, then the velocity will change, or will not remain constant. This is illustrated in Figure 8.3.4.

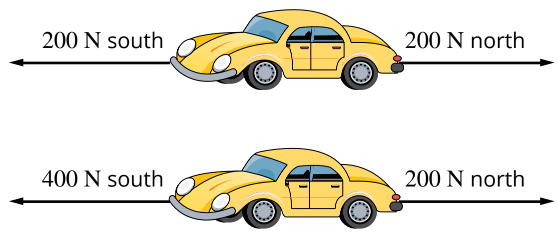


FIGURE 8.3.4 The forces on the top car are balanced, and so it will maintain a constant velocity. The forces on the bottom car are unbalanced: it has a net force in the forwards direction, so its velocity will change (it will speed up).

PHYSICSFILE

The effects of forces

Applying a force can cause an object to speed up, to slow down, to start moving, to stop moving, or to change its direction. The effect depends on the direction of the force in relation to the direction of the velocity vector of the object experiencing the force. The effect of external forces is summarised in Table 8.3.1 below.

TABLE 8.3.1 The effect of the application of a force, depending on the relationship between the direction of the force and the velocity.

Relationship between velocity and force	Effect of force
force applied to object at rest	object starts moving
force in same direction as velocity	magnitude of velocity increases (object speeds up)
force in opposite direction to velocity	magnitude of velocity decreases (object slows down)
force perpendicular to velocity	direction of velocity changes (object turns)

In all cases, the effect of a force is to change the velocity of an object, whether it is the magnitude of the velocity, the direction of the velocity, or both that changes.

FIGURE 8.3.3 The upwards thrust force on a rocket has the effect of speeding it up, whereas the upwards drag force on a parachute has the effect of slowing the descent of the parascenders.

The term 'external', in relation to forces, implies that the forces are not internal. When forces are internal, they will have no effect on the motion of the object. For example, if you are sitting in a car and push forwards on the steering wheel then the car will not move forwards due to this force. In order for you to push forwards on the steering wheel, you must push backwards on the seat. Both the steering wheel and the seat are attached to the car, therefore there are two forces acting on the car that are equal and in the opposite direction to each other, as shown in Figure 8.3.5. All internal forces must result in balanced forces on the object and therefore they will not change the velocity of the object.

FIGURE 8.3.5 This driver is applying internal forces on a car. These internal forces will balance and cancel each other.

Stating Newton's first law in different ways

Ludwig Wittgenstein, an Austrian-British philosopher, suggested that 'understanding means seeing that the same thing said different ways is the same thing'. To truly understand Newton's first law, you should be able to state it in different ways yet still recognise it as being consistent with Newton's first law.

All of these statements are consistent with Newton's first law:

- An object will maintain a constant velocity unless an unbalanced, external force acts on it.
- An object will continue with its motion unless an unbalanced, external force is applied.
- An object will not continue with its velocity if an unbalanced, external force is applied.
- A body will either remain at rest or continue with constant speed in a straight line (i.e. constant velocity) unless it is acted on by a net force.
- If an unbalanced, external force is applied, then an object's velocity will change.
- If a net force is applied, then the object's velocity will change.
- If no net force is applied, the object will not accelerate.
- If a net force is applied, the object will accelerate.
- Net forces cause acceleration.
- No force, no acceleration.
- Constant velocity means no net force is applied.

PHYSICS IN ACTION

Terminal velocity

In Chapter 7 it was stated that, in the absence of air resistance, all objects accelerate towards the surface of Earth at a constant rate of $9.8\,m\,s^{-2}$. However, air resistance increases as the speed of an object increases, and it becomes very significant when objects are moving at high speeds. Newton's first law can be used to explain how air resistance causes a skydiver to experience terminal velocity. As the skydiver begins falling towards Earth the only external force is the weight force. The weight force causes the skydiver to accelerate at $9.8\,m\,s^{-2}$. As the skydiver gets faster, air resistance pushes upwards. This reduces the magnitude of the net force, which decreases the skydiver's acceleration. Eventually the air resistance becomes so great that it exactly balances the weight force. According to Newton's first law, the skydiver maintains a constant velocity when all the external forces are balanced: this velocity is terminal velocity. Terminal velocity will be different for different objects because air resistance depends on the object's shape. For example, terminal velocity for a sheet of paper is much less than for a skydiver.

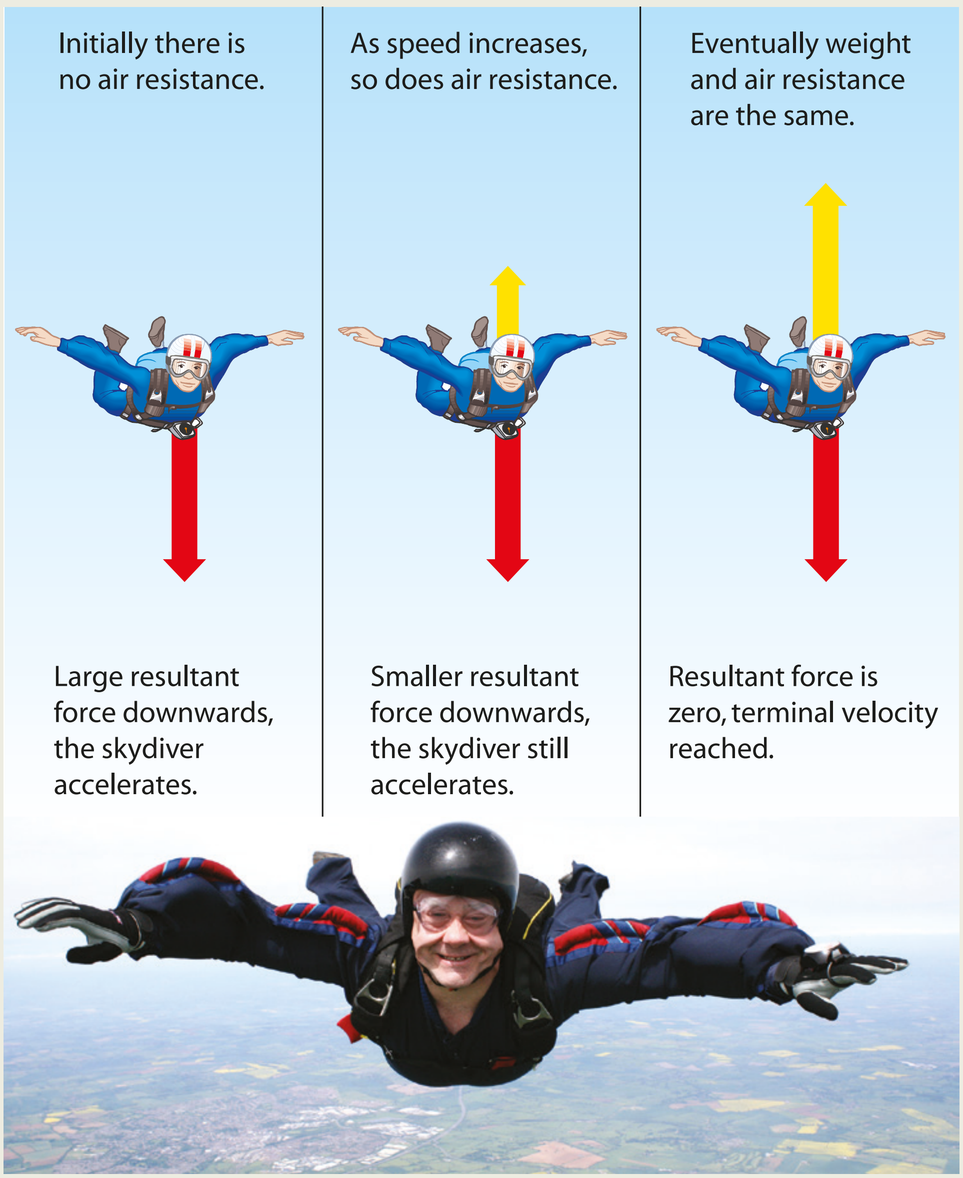


FIGURE 8.3.6 When a skydiver reaches terminal velocity, weight force is in balance with air resistance.

INERTIA

Inertia is considered to be the resistance to a change in motion of an object. It is related to the mass of the object. As the mass of the object increases, the inertia increases and therefore:

- it becomes harder to start it moving if it is stationary, or
- it becomes harder to stop it moving, or
- it becomes harder to change its direction of motion.

You can experience the effect of inertia when you push a trolley in a supermarket. If the trolley is empty, it is relatively easy to start pushing it, or to pull it to a stop when it is already moving. It is also easy to turn a corner. If you fill the trolley with heavy groceries, you notice that it becomes more difficult—that is, it requires more force—to make the trolley start moving when it is at rest, and it becomes more difficult to pull it to a stop if it is already moving. It also requires more force to change the trolley's direction. This concept was explored in sections 8.1 and 8.2 with discussions on momentum. You learned that the greater an object's mass or velocity, the greater its momentum, and consequently the harder it is to slow it down. This is a direct relationship to its inertia.

It is important to note that the effect of inertia is independent of gravity. Since inertia depends on mass, and weight force due to gravity also depends on mass, it is a common misunderstanding to think that the effects of inertia only apply in the presence of gravity. However, even in space it would be just as difficult to change the state of motion of the trolley, as described above, as it is in the supermarket aisle.

Newton's first law and inertia

The connection between Newton's first law and inertia is very close. Due to inertia, an object will continue with its motion unless a net force acts on the object.

You experience the connection between Newton's first law and inertia if you are standing on public transport. Imagine standing on a train that is initially at rest and then starts moving forwards. If you are not holding on to anything, you may stumble backwards as though you have been pushed backwards. However, you have not been pushed backwards; the train has started moving forwards and, since you have inertia, your mass resists the change in motion. According to Newton's first law, your body is simply maintaining its original state of being motionless until an unbalanced force acts to accelerate it. When the train later comes to a sudden stop, your body again resists the change by continuing to move forwards until an unbalanced force acts to bring it to a stop.

OBSERVING NEWTON'S FIRST LAW

When an object is in motion—for example, a pen sliding across a table—it will eventually stop. It may not seem obvious, but this is a very good example of Newton's first law. The motion does not continue; therefore a net force must be acting. In this case, however, the force is not an obvious one. Confusion sometimes arises if the force due to friction is overlooked. Friction is a force that always acts in the opposite direction to the motion of objects. Air resistance is also a force that is often overlooked, as is the force due to gravity. By ignoring the effect of these important forces, it can be easy to come to the incorrect conclusion that the natural state of any object is to be at rest. By considering all the external forces acting on an object, it becomes clear that the natural state of any object is to maintain whatever velocity it currently has.

EXTENSION

Frictional forces

Friction is a force that opposes movement. Suppose you want to push your textbook along the table. As you start to push the book, you find that the book does not move at first. You then increase the force that you apply. Suddenly, at a certain critical value, the book starts to move.

There is a maximum frictional force that resists the start of the slide. This force is called the static friction force, F_s. Once the book begins to slide, a much lower force than F_s is needed to keep the book moving. This force is called the kinetic friction force and is represented by F_k. The graph in Figure 8.3.7 shows how the force required to move an object changes as static friction is overcome.

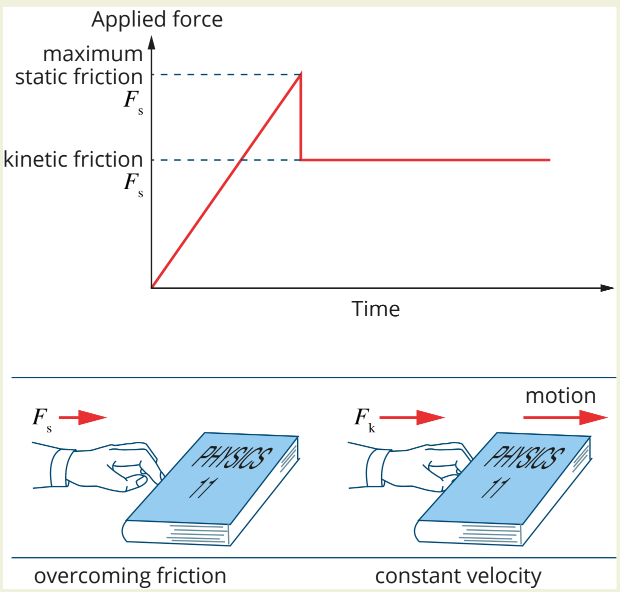


FIGURE 8.3.7 To get things moving, the static friction between an object and the surface must be overcome. This requires a larger force than is needed to maintain constant velocity.

This phenomenon can be understood when you consider that even the smoothest surfaces are quite jagged at the microscopic level. When the book is resting on the table, the jagged points of its bottom surface have settled into the valleys of the surface of the table and this helps to resist attempts to slide the book. Once the book is moving, the surfaces do not have any time to settle into each other, so less force is required to keep the book moving.

Another fact that helps to explain friction arises from the forces of attraction between the atoms and molecules of the two different surfaces that are in contact. These surfaces produce weak bonding between their respective particles, so before one surface can move across the other, these bonds must be broken. This extra effort adds to the static friction force. Once there is relative motion between the surfaces, the bonds cannot re-form.

In everyday life, there are situations in which friction is desirable (e.g. walking) and others in which it is a definite problem. Consider the moving parts within the engine of a car. Friction can rob an engine of its fuel economy and cause it to wear out. Special oils and lubricants are introduced in order to prevent moving metal surfaces from touching. If the moving surfaces actually moved over each other, they would quickly wear, producing metal filings that could damage the engine. Instead, metal surfaces are separated by a thin layer of oil. The oils are chosen on the basis of their viscosity (thickness). For example, low viscosity oils can be used in the engine, while heavier oils are needed in the gearbox and differential of the car where greater forces are applied to the moving parts. On the other hand, the lack of friction enables magnetic levitation trains to reach very high speeds, as shown in Figure 8.3.8.

FIGURE 8.3.8 This magnetic levitation train in China rides 1 cm above the track, so the frictional forces are negligible. The train is propelled by a magnetic force to a cruising speed of about 430 km h^{-1}.

At other times, having friction is essential. If there is any ice on the road when driving to snowfields, drivers are required to fit chains to their cars. When driving over a patch of ice, the chains break through the ice and the car is able to grip the road. Similarly, friction is definitely required within the car's brakes when the driver wants to slow down. In fact, modern brake pads are specially designed to maximise the friction between them and the brake drum or disc.

PHYSICS IN ACTION

Galileo's law of inertia

Galileo Galilei (Figure 8.3.9) was born into an academic family in Pisa, Italy, in 1564. He made significant contributions to physics, mathematics and scientific method through intellectual rigour and the quality of his experimental design. But, more than this, Galileo helped to change the way the universe is viewed.

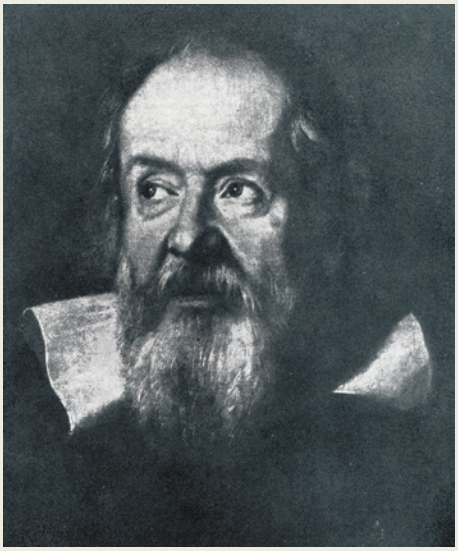


FIGURE 8.3.9 Galileo Galilei.

Galileo's most significant contributions were in astronomy. Through his development of the refracting telescope he discovered sunspots, lunar mountains and valleys, the four largest moons of Jupiter (now called the Galilean Moons, one of which is shown in Figure 8.3.10) and the phases of Venus. In mechanics, he demonstrated that projectiles moved with a parabolic path and that different masses fall at the same rate (the law of falling bodies).

FIGURE 8.3.10 The major moons of Jupiter, including Io, are known as the Galilean Moons.

These developments were important because they changed the framework within which mechanics was understood. This framework had been in place since Aristotle had constructed it in the fourth century BCE. Aristotle's thesis was based on the observation that a moving body's natural state is at rest, and the object will come to rest unless a force is applied. However, Galileo's experiments led him to believe that the natural state of an object is not at rest. He suggested that objects maintained their state of motion. He called this tendency inertia.

By the sixteenth century, the work of the Greek philosophers had become entrenched, being widely supported in the universities and at a political level. In Italy at that time, government was controlled by the Roman Catholic Church. One might think that Galileo would have won praise from his peers for making such progress, but so ingrained and supported was the Aristotelian view that Galileo actually lost his job as a professor of mathematics in Pisa in 1592.

Galileo was not, however, without supporters, and he was able to move from Pisa to Padua, where he continued teaching mathematics. In Padua, Galileo began to use measurements from carefully constructed experiments to strengthen his ideas. He entered into vigorous debate in which his ideas (founded on observation) were pitted against the philosophy of the past and the politics of the day. The most divisive debate involved the motion of the planets. The ancient Greek view, formalised by Ptolemy in the second century CE, was that Earth was at the centre of the solar system and that the planets, the Moon and the Sun were in orbit around it. This view was taught by the Church and was also supported by common sense.

In 1630, Galileo published a book in which he debated the Ptolemaic view and the new Sun-centred model proposed by Copernicus. On the basis of his own observations, Galileo supported the Copernican view of the universe. Although the book had been passed by the censors of the day, Galileo was summoned to Rome to face the Inquisition for heresy (opposition to the Church). The finding went against Galileo and all copies of his book had to be burned. He was sentenced to permanent house arrest for the rest of his life.

Galileo died in 1642 in a village near Florence. He had become an influential thinker across Europe, and the scientific revolution he had helped to start accelerated in the freer Protestant countries in northern Europe. For its part, the Roman Catholic Church under Pope John Paul II in 1979 began an investigation into Galileo's trial, and in 1992 a papal commission reversed the Church's condemnation of him.

8.3 Review

SUMMARY

- A force is a push or a pull. Some forces act on contact while others can act at a distance.
- Force is a vector quantity whose SI unit is the newton (N).
- Newton's first law can be written in many ways:
  - An object will continue with its velocity unless an unbalanced force causes the velocity to change.
  - Net forces cause acceleration.
- Inertia is the tendency of an object to resist changes in motion.
- Inertia is related to mass; an object with a large mass will have a large inertia.

KEY QUESTIONS

1 A student observes a box sliding across a surface and slowing down to a stop. From this observation what can the student conclude about the forces acting on the box?

2 A car changes its direction as it turns a bend in the road while maintaining its speed of $16\,m\,s^{-1}$. From this, what can you conclude?

3 A bowling ball rolls along a smooth wooden floor at constant velocity. Ignoring the effects of friction and air resistance, which of the following options, relating to the force acting on the ball, is correct?

A There must be a net force acting forwards to maintain the velocity of the ball.
B There must not be an unbalanced force acting on the ball.
C The forwards force acting on the ball must be balanced by the friction that opposes the motion.
D More information is needed.

4 If a person is standing up in a moving bus that stops suddenly, the person will tend to fall forwards. Has a force acted to push the person forwards? Use Newton's first law to explain what is happening.

5 What horizontal force has to be applied to a wheelie bin if it is to be wheeled to the street on a horizontal path against a frictional force of 20 N at a constant $1.5\,m\,s^{-1}$?

6 A young boy is using a horizontal rope to pull his billycart at a constant velocity. A frictional force of 25 N also acts on the billycart.

a What force must the boy apply to the rope?
b The boy's father then attaches a longer rope to the cart because the short rope is uncomfortable to use. The rope now makes an angle of 30° to the horizontal. What is the horizontal component of the force that the boy needs to apply in order to move the cart with constant velocity?
c What is the tension force acting along the rope that the boy must supply?

7 Passengers on commercial flights are required to be seated and have their seatbelts done up when their plane is coming in to land. What would happen to a person who was standing in the aisle as the plane travelled along the runway during landing?

8 Consider the following situations, and name the force that causes each object to travel along a path which is not a straight line.

a The Earth moves in a circle around the Sun with constant speed.
b An electron orbits the nucleus with constant speed.
c A cyclist turns a corner at constant speed.
d An athlete swings a hammer in a circle with constant speed.

9 A magician performs a trick in which a cloth is pulled quickly from under a glass filled with water without causing the glass to fall over or the water to spill out.

a Explain the physics principles underlying this trick.
b Does using a full glass make the trick easier or more difficult? Explain.

10 Which of these objects would find it most difficult to come to a stop: a cyclist travelling at $50\,km\,h^{-1}$, a car travelling at $50\,km\,h^{-1}$ or a fully laden semitrailer travelling at $50\,km\,h^{-1}$? Explain your answer.

11 When flying at constant speed at a constant altitude, a light aircraft has a weight of 50 kN down, and the thrust produced by its engines is 12 kN to the east. What is the lift force required by the wings of the plane, and how large is the drag force that is acting?

8.4 Newton's second law

Newton's second law makes the quantitative connection between force, mass and acceleration.

Newton's second law helps to resolve the misconception that many people have about the time taken for objects of different mass to fall to the ground. Many mistakenly believe that heavy objects will fall faster than lighter objects. Once again, air resistance acts to complicate the matter and results in the observation of different times for different masses to fall the same distance. However, even when air is removed, the misconception that larger masses fall faster than lighter masses still persists.

Figure 8.4.1 depicts a famous experiment, mentioned in Section 8.4 of this text, which shows that objects fall together when the effect of air resistance is removed. A web search for 'hammer and feather on the Moon' will enable you to view a video of David Scott's 1971 experiment. Although the images are quite poor, you should be able to see both objects accelerating at the same rate.

FIGURE 8.4.1 An artist's image of the famous hammer and feather experiment conducted on the Moon.

NEWTON'S SECOND LAW

Newton's second law of motion states that:

> The acceleration of an object is directly proportional to the net force on the object and inversely proportional to the mass of the object:
>
> $$a = \frac{F_{net}}{m}$$
>
> where a is the acceleration of an object (in $m\,s^{-2}$)
>
> F_{net} is the force applied to the object (in N)
>
> m is the mass of the object (in kg).

The above equation is also commonly written as $F_{net} = ma$. F_{net} is the unbalanced force applied to the object.

The unit of force is the units of mass and acceleration combined, or $kg\,m\,s^{-2}$. This unit was renamed the newton (N) in honour of Sir Isaac Newton. By definition, 1 **newton** is the force needed to accelerate a mass of 1 kg at $1\,m\,s^{-2}$.

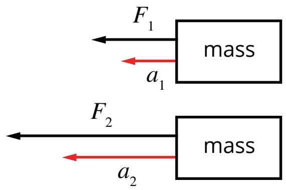


FIGURE 8.4.2 Given the same mass, a larger force will result in a larger acceleration. If the force is doubled, then the acceleration is also doubled.

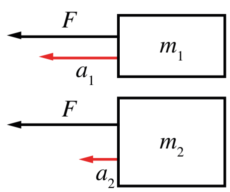


FIGURE 8.4.3 Given the same force, a larger mass will result in a lower acceleration. If the mass is doubled, then the acceleration is halved.

One of the implications of Newton's second law is that, for a given mass, a greater acceleration is achieved by applying a greater force. This is shown in Figure 8.4.2. Doubling the applied force will double the acceleration of the object. In other words, acceleration is proportional to the net force applied.

Notice also in Figure 8.4.2 that the acceleration of the object is in the same direction as the net force applied to it.

Newton's second law also explains how acceleration is affected by the mass of an object. For a given force, the acceleration of an object will decrease with increased mass. In other words, acceleration is inversely proportional to the mass of an object. This is shown in Figure 8.4.3.

Furthermore, as suggested in Section 8.1, Newton's second law can be used to show that force is equal to the rate of change of momentum.

The following derivation will show how Newton's second law relates to momentum. It shows that the net force, F_{net}, is equal to the change in momentum, Δp, divided by the period of time, Δt, which is the rate of change of momentum:

$$\begin{aligned} F_{net} &= ma \\ &= m\frac{(v-u)}{\Delta t} \\ &= \frac{mv-mu}{\Delta t} \\ &= \frac{\Delta p}{\Delta t} \end{aligned}$$

This means that changes in momentum are caused by the action of a net force. This should seem reasonable, since you have now established that a net force causes the acceleration of an object, and hence a change of its velocity, which therefore changes its momentum.

PHYSICSFILE

Maximising acceleration

Dragster race cars are designed to achieve the maximum possible acceleration in order to win a race in a straight line over a relatively short distance. According to Newton's second law, acceleration is increased by increasing the applied force and by reducing the mass of the object. For this reason, dragster race cars are designed with very powerful engines that produce an enormous forwards force and an aerodynamic shape to minimise air resistance. There is not much else to the car, so this helps to minimise the mass.

Newton's second law also helps you understand why a motorcycle can accelerate from the lights at a greater rate than a car or a truck. While the engines in a car or truck are usually more powerful than a motorcycle engine, the motorcycle has much less mass, which allows for greater acceleration.

FIGURE 8.4.4 The aerodynamic design of motorcycles and their lower mass enables them to accelerate faster than cars and trucks.

Observing Newton's second law

The equation $F_{net} = ma$ enables you to calculate the force that causes a mass to accelerate. Mass is something that is easily experienced. You can measure the mass of an object on a balance. You can hold two masses in your hands and feel their effect. Similarly, you can observe acceleration, for example when a car accelerates away from the traffic lights. Force, on the other hand, is not something you can see. However, you can see the *effect* of a force.

Worked example 8.4.1

CALCULATING THE FORCE THAT CAUSES AN ACCELERATION

Calculate the net force causing a 5.50 kg mass to accelerate at 3.75 m s^{-2} west.	
Thinking	**Working**
Ensure that the variables are in their standard units.	$m = 5.50$ kg $a = 3.75$ m s^{-2} west
Apply the equation for force from Newton's second law.	$F_{net} = ma$ $= 5.50 \times 3.75$ $= 20.6$ N
Give the direction of the net force, which is always the same as the direction of the acceleration.	$F_{net} = 20.6$ N west

Worked example: Try yourself 8.4.1

CALCULATING THE FORCE THAT CAUSES AN ACCELERATION

Calculate the net force causing a 75.8 kg runner to accelerate at 4.05 m s^{-2} south.

The first equation for uniform acceleration, which is discussed in Chapter 7 on page XX, can be combined with Newton's second law to calculate changes in time or velocity.

The first equation for uniform acceleration is:

$v = u + at$

This can be rearranged to give:

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

Combining this with $F_{net} = ma$ gives:

$$F_{net} = m\left(\frac{v-u}{t}\right)$$

Worked example 8.4.2

CALCULATING THE FINAL VELOCITY OF AN ACCELERATING MASS

Calculate the final velocity of a 225 kg scooter that accelerates for 2.00 s from rest due to a force of 2430 N north.	
Thinking	**Working**
Ensure that the variables are in their standard units.	$m = 225$ kg $t = 2.00$ s $u = 0$ m s^{-1} $F_{net} = 2430$ N north
Apply a variation of the equation for force from Newton's second law.	$F_{net} = m\left(\frac{v-u}{t}\right)$ $(v-u) = \frac{F_{net}t}{m}$ $v = \frac{F_{net}t}{m} + u$ $= \frac{2430 \times 2.00}{225} + 0$ $= 21.6$ m s^{-1}
Give the direction of the final velocity as being the same as the direction of the force.	$v = 21.6$ m s^{-1} north

Worked example: Try yourself 8.4.2

CALCULATING THE FINAL VELOCITY OF AN ACCELERATING MASS

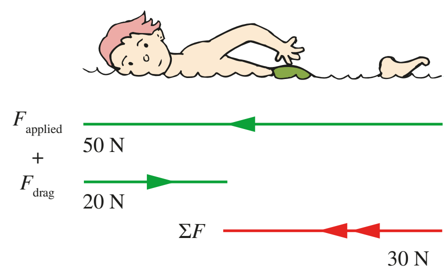
Calculate the final velocity of a 307 g fish that accelerates for 5.20 s from rest due to a force of 0.250 N left.

Forces do not always act alone. Mostly, more than one force will act on an object at any time. The overall effect of the forces depends on the direction of each of the forces. For example, some forces act together and some may oppose each other. When using Newton's second law, it is important to use the net, or resultant, force in the calculation. As forces are vectors, they can be added or combined using the techniques discussed in Chapter 6. Consider the following worked examples.

Worked example 8.4.3

CALCULATING THE ACCELERATION OF AN OBJECT WITH MORE THAN ONE FORCE ACTING ON IT

A swimmer whose mass is 75 kg applies a force of 50 N as he starts a lap. The water opposes his efforts to accelerate with a drag force of 20 N. What is his initial acceleration?

Thinking	Working
Determine the individual forces acting on the swimmer, and apply the vector sign convention.	F_1 = 50 N forwards $= 50$ N F_2 = 20 N backwards $= -20$ N
Determine the net force acting on the swimmer.	$F_{net} = F_1 + F_2$ $= 50 + (-20)$ $= +30$ N or 30 N forwards 
Use Newton's second law to determine the acceleration.	$a = \frac{F_{net}}{m}$ $= \frac{30}{75}$ $= 0.40$ m s^{-2} forwards

Worked example: Try yourself 8.4.3

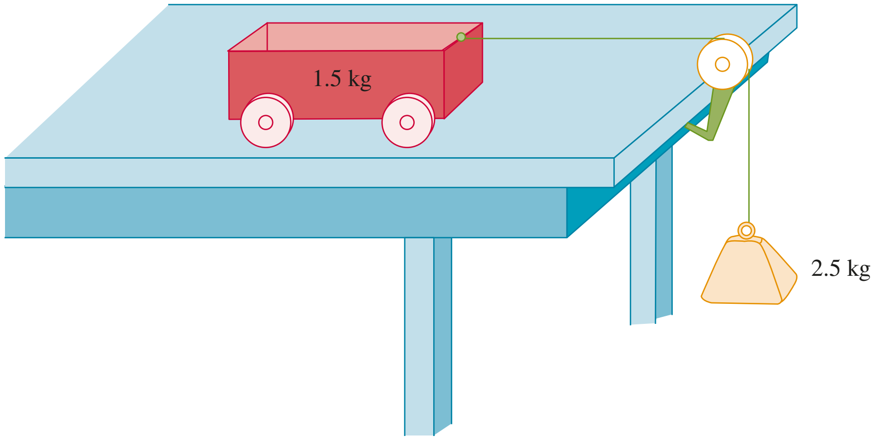
CALCULATING THE ACCELERATION OF AN OBJECT WITH MORE THAN ONE FORCE ACTING ON IT

A car with a mass of 900 kg applies a driving force of 3000 N as it starts moving. Friction and air resistance oppose the motion of the car with a force of 750 N. What is the car's initial acceleration?

Worked example 8.4.4

CALCULATING THE ACCELERATION OF A CONNECTED BODY

A 1.5 kg trolley cart is connected by a cord to a 2.5 kg mass as shown. The cord is placed over a pulley and the mass is allowed to fall under the influence of gravity.



a Assuming that the cart can move over the table without friction, determine the acceleration of the cart.	
Thinking	**Working**
Recognise that the cart and the falling mass are connected, and determine a sign convention for the motion.	As the mass falls, the cart will move to the right. Therefore, both downwards movement of the mass and rightwards movement of the cart will be considered positive motion.
Write down the data that is given. Apply the sign convention to vectors.	$m_1 = 2.5\,\text{kg}$ $m_2 = 1.5\,\text{kg}$ $g = 9.8\,\text{m s}^{-2}$ down $= +9.8\,\text{m s}^{-2}$
Determine the forces acting on the system.	The only force acting on the combined system of the cart and mass is the weight of the falling mass. $F_{\text{net}} = F_g$ $= m_1 g$ $= 2.5 \times 9.8$ $= 24.5\,\text{N}$ in the positive direction
Calculate the mass being accelerated.	This net force has to accelerate not only the cart but also the falling mass. $m_1 + m_2 = 2.5 + 1.5$ $= 4.0\,\text{kg}$
Use Newton's second law to determine acceleration.	$a = \dfrac{F_{\text{net}}}{m}$ $= \dfrac{24.5}{4.0}$ $= 6.1\,\text{m s}^{-2}$ to the right

b If a frictional force of 8.5 N acts against the cart, what is the acceleration now?	
Thinking	**Working**
Write down the data that is given. Apply the sign convention to vectors.	$m_1 = 2.5\,\text{kg}$ $m_2 = 1.5\,\text{kg}$ $g = 9.8\,\text{m s}^{-2}$ down $= +9.8\,\text{m s}^{-2}$ $F_{fr} = 8.5\,\text{N}$ left $= -8.5\,\text{N}$
Determine the forces acting on the system.	There are now two forces acting on the combined system of the cart and mass: the weight of the falling mass and friction. $F_{net} = F_g + F_{fr}$ $= 24.5 + (-8.5)$ $= 16.0\,\text{N}$ $= 16.0\,\text{N}$ in the positive direction
Use Newton's second law to determine acceleration.	$a = \frac{F_{net}}{m}$ $= \frac{16.0}{4.0}$ $= 4.0\,\text{m s}^{-2}$ to the right

Worked example: Try yourself 8.4.4

CALCULATING THE ACCELERATION OF A CONNECTED BODY

A 0.6 kg trolley cart is connected by a cord to a 1.5 kg mass. The cord is placed over a pulley and the mass is allowed to fall under the influence of gravity.

a Assuming that the cart can move over the table unhindered by friction, determine the acceleration of the cart.

b If a frictional force of 4.2 N acts against the cart, what is the acceleration now?

THE FEATHER AND HAMMER EXPERIMENT

Why the experiment works on the Moon

When two objects with different mass fall under the influence of the force due to gravity in the absence of air resistance, they will both fall at the same rate. That is, their accelerations will be the same. They will cover the same displacement in the same time and will hit the ground at the same time if dropped from the same height. This experiment works on the Moon because there is no atmosphere and, therefore, no air resistance.

If you understand that all objects accelerate due to gravity at the same rate in a vacuum, the common misconception is that the force due to gravity is the same on all objects. This is not true. In fact, the force due to gravity is larger on larger objects and smaller on smaller objects. You mustn't forget that the objects have different masses. A larger mass experiences a greater force due to gravity (weight) than a smaller mass, but it also has more inertia so it requires that greater force in order to achieve the same acceleration. Refer to Figure 8.4.5 to see how this works.

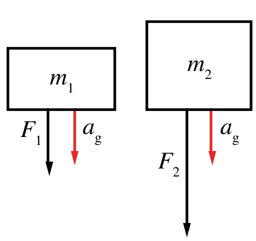


FIGURE 8.4.5 Given the same acceleration (a_g), a larger mass (m_2) must have a larger force (F_2) acting on it. If the mass is doubled, then the force is doubled.

If you consider the relationship between acceleration and mass, and the relationship between acceleration and force, then:

$$a = \frac{F}{m}$$

If m_2 is ten times the mass of m_1, then the force due to gravity on m_2 is ten times the force due to gravity on m_1. Consider the acceleration of both masses:

$$a_1 = \frac{F_1}{m_1}$$

If $F_2 = 10F_1$ and $m_2 = 10m_1$, then:

$$\begin{aligned} a_2 &= \frac{F_2}{m_2} \\ &= \frac{10F_1}{10m_1} \\ &= \frac{F_1}{m_1} \\ &= a_1 \end{aligned}$$

This proof shows that the ratio of force to mass is equal for all combinations of force and mass under the same effects due to gravity. This proves that all masses will experience the same acceleration if air resistance is removed.

Why the experiment does not work on Earth

When you see a feather floating down through the air, you know that it is accelerating at a rate far less than a hammer falling from the same height. From the previous section, you will know that the hammer and the feather have forces due to gravity acting on them that are proportional to their mass. They do not fall at the same rate because of the force of air resistance. Remember, Newton's second law says the acceleration is proportional to the *net* force acting on an object, which means you must consider all the forces acting on an object to determine the acceleration.

Air resistance is a force that results from air molecules colliding with the object. The faster the object moves, the greater the air resistance. In addition, the greater the surface area perpendicular to the direction of motion, the greater the air resistance. This force, which acts in the opposite direction to the motion of an object, is significant when compared with the weight of the feather, but insignificant when compared with the weight of the hammer. As a result, this force has a noticeable effect on the feather's acceleration but makes no noticeable difference to the hammer's acceleration.

In Figure 8.4.6, the force of air resistance is denoted as F_{AR} and is the same size on both objects. The difference between the two objects is the downwards weight force (due to gravity).

Figure 8.4.6 also shows that the acceleration of the thinner object is much less than the acceleration of the thicker object. This is the observation that often causes misconceptions.

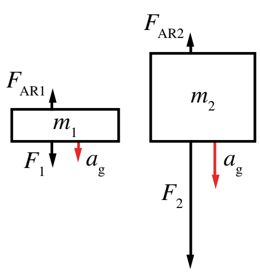


FIGURE 8.4.6 The effect of air resistance on an object depends on the surface area perpendicular to the motion and the speed of the object.

8.4 Review

SUMMARY

- Newton's second law states: The acceleration of an object is directly proportional to the net force on the object and inversely proportional to the mass of the object.
- Force can be calculated using the following formulas:

  $F_{net} = ma$
- This can be rewritten as:

  $$F_{net} = m\left(\frac{v-u}{t}\right) \text{ or } F_{net} = \frac{\Delta p}{\Delta t}$$
- Force is difficult to perceive when it acts on objects, but we can perceive mass and acceleration.
- Different forces due to gravity act on different masses to cause the same acceleration.
- Air resistance is a force that acts to decrease the acceleration of objects moving through air.

KEY QUESTIONS

Use $g = 9.8\,\text{m}\,\text{s}^{-2}$ when answering the following questions.

1 Calculate the acceleration of a 23.9 kg mass when a net force of 158 N north acts on it.

2 Calculate the mass of an object if it accelerates at $9.20\,\text{m}\,\text{s}^{-2}$ east when a net force of 352 N east acts on it.

3 Calculate the final velocity of a stationary 55.9 kg mass when a net force of 56.8 N north acts on it for 3.50 seconds.

4 Calculate the acceleration of a 45.0 kg mass that has a net force of 441 N acting on it due to gravity.

5 Calculate the acceleration of a 90.0 kg mass that has a net force of 882 N acting on it due to gravity.

6 Calculate the final velocity of a 60.0 kg mass moving at $2.67\,\text{m}\,\text{s}^{-1}$ east, when a net force of 45.5 N west acts on it for 2.80 seconds.

7 Calculate the acceleration of a 60.9 g golf ball when a net force of 95.0 N south acts on it.

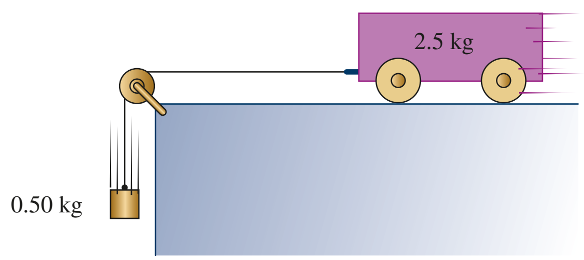
8 Calculate the mass of a train if it accelerates at $7.20\,\text{m}\,\text{s}^{-2}$ north when a net force of 565 000 N north acts on it. Give your answer to three significant figures.

9 Calculate the final velocity of a stationary 3.00 g marble when a net force of 0.0823 N north acts on it for 0.0105 seconds.

10 Mary is paddling a canoe. The paddles are providing a constant driving force of 45 N south and the drag forces total 25 N north. The mass of the canoe is 15 kg and Mary has a mass of 50 kg.

- **a** What is Mary's mass?
- **b** Calculate Mary's weight.
- **c** Find the net horizontal force acting on the canoe.
- **d** Calculate the magnitude of the acceleration of the canoe.

11 A 0.50 kg metal block is attached by a piece of string to a dynamics cart, as shown below. The block is allowed to fall from rest, dragging the cart along. The mass of the cart is 2.5 kg.



- **a** If friction is ignored, what is the acceleration of the block as it falls?
- **b** How fast will the block be travelling after 0.50 s?
- **c** If a frictional force of 4.3 N acts on the cart, what is its acceleration?

12 An empty truck of mass 2000 kg has a top acceleration of $2.0\,\text{m}\,\text{s}^{-2}$. The mass of one box is 300 kg. How many boxes would be loaded if the truck's top acceleration decreased to $1.25\,\text{m}\,\text{s}^{-2}$.

13 The thrust force of a rocket with a mass of 50 000 kg is 1 000 000 N. Neglecting air resistance, calculate its acceleration.

8.5 Newton's third law

Newton's first two laws of motion describe the motion of an object resulting from the forces that act on that object. Newton's third law of action and reaction is easily stated and seems to be widely known by students, but it is often misunderstood and misused. It is a very important law in physics as it assists with the understanding of the origin and nature of forces. Newton's third law is explored in detail in this section.

NEWTON'S THIRD LAW

Newton realised that all forces exist in pairs and that each force in the pair acts on a different object. Look at Figure 8.5.1, which shows a hammer hitting a nail on the head. Both the hammer and the nail experience forces during this interaction. The nail experiences a downwards force as the hammer hits it. When the nail is hit it moves a distance into the wood. As it hits the nail, the hammer experiences an upwards force that causes the hammer to stop. These forces are known as an action-reaction pair and are shown in Figure 8.5.2.

FIGURE 8.5.1 A hammer hitting a nail is a good example of an action–reaction pair and Newton's third law.

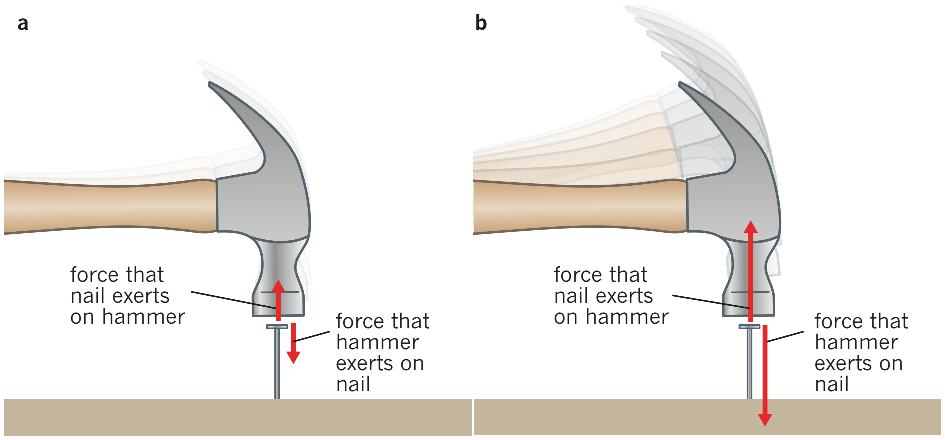


FIGURE 8.5.2 (a) As the hammer gently taps the nail, both the hammer and the nail experience small forces. (b) When the hammer smashes into the nail, both the hammer and the nail experience large forces.

It is important to note that, regardless of whether the hammer exerts a small or a large force on the nail, the nail will exert exactly the same-size force back on the hammer.

Newton's third law states:

> For every action (force), there is an equal and opposite reaction (force).

This means that when object A exerts a force, F, on object B, object B will exert an equal and opposite force on object A. It is important to recognise that the action force and the reaction force in Newton's third law act on different objects and so should never be added together; their effect will only be on the object on which they act. Newton's third law applies not only to forces between objects which are in direct contact, but also to non-contact forces, such as the force due to gravity between objects.

The main misconception that arises when considering Newton's third law is the belief that, if a large mass collides with a smaller mass, then the larger object exerts a larger force and the smaller object exerts a smaller force. This is not true. If you witnessed the collision between the car and the bus in Figure 8.5.3c, you would see the car undergoing a large deceleration while the bus undergoes only a small acceleration.

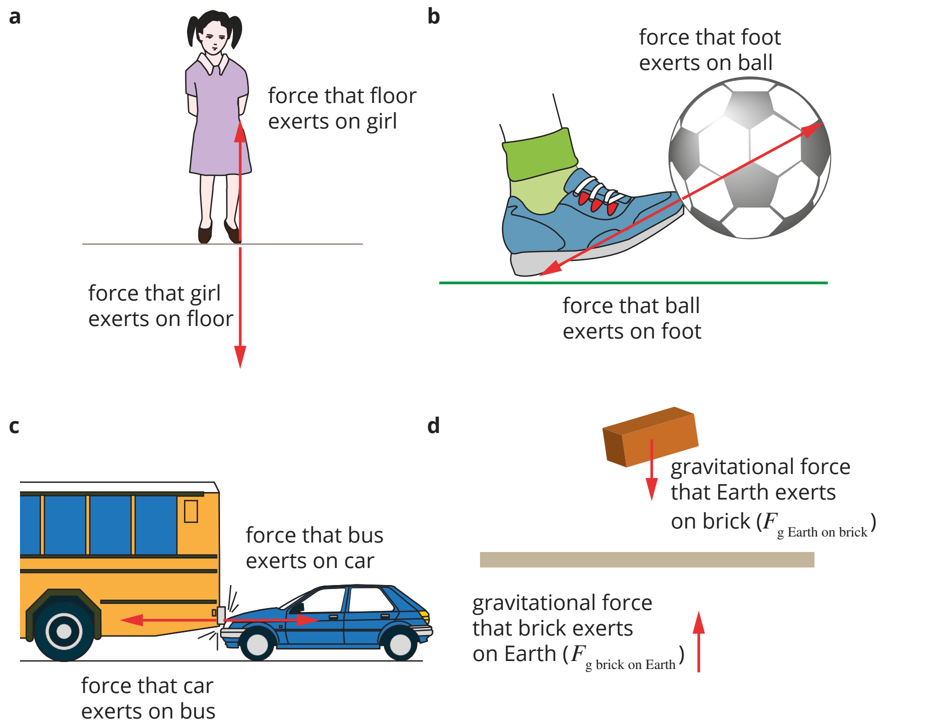


FIGURE 8.5.3 Some action—reaction pairs.

From Newton's second law, you know that the same force acting on a larger mass will result in a smaller acceleration. This is the effect seen in the situation of the car colliding with the bus. Because of the car's small mass, the force acting on the car will cause the car to undergo a large deceleration. The occupants may be seriously injured as a result of this. The force acting on the bus is equal in size, but is acting on a much larger mass. As a result, the bus will have a relatively small acceleration and the occupants will not be as seriously affected.

Identifying the action and reaction forces

When analysing a situation to determine the action and reaction forces according to Newton's third law, it is helpful to be able to label the force vectors systematically. A good strategy for labelling force vectors is to use the capital letter F to represent the force and then to include a subscript consisting of the word 'on' and the thing on which the force is acting, and then the word 'by' and the thing that is applying the force.

The equal and opposite force is then labelled with a capital F and a subscript with the objects in reverse. For example, the action and reaction force vector arrows shown in Figure 8.5.3 can be labelled as shown in Table 8.5.1.

TABLE 8.5.1 Labels of action and reaction force vectors in Figure 8.5.3.

	Action vector	Reaction vector
a	$F_{\text{on floor by girl}}$	$F_{\text{on girl by floor}}$
b	$F_{\text{on ball by foot}}$	$F_{\text{on foot by ball}}$
c	$F_{\text{on bus by car}}$	$F_{\text{on car by bus}}$
d	$F_{\text{on Earth by brick}}$	$F_{\text{on brick by Earth}}$

It should be noted that it does not matter which force is considered the action force and which is considered the reaction force. They are always equal in magnitude and opposite in direction.

PHYSICSFILE

Combining Newton's second and third laws in the classroom

You can easily observe the effect of Newton's second and third law in the classroom if you have two dynamics carts with wheels that are free to roll on a smooth surface (such as a bench or desk). If the two carts are placed in contact with each other, and the plunger is activated on *one* of the carts, you will observe that *both* carts roll backwards. This is because of the action and reaction force pair described by Newton's third law. If the two carts have similar masses, you will observe that they accelerate apart at a similar rate. If one cart is heavier than the other, you will observe the lighter cart accelerates at a greater rate. This is because the forces acting on both carts are equal in magnitude and so, according to Newton's second law, the smaller mass will experience a greater acceleration.

Worked example 8.5.1

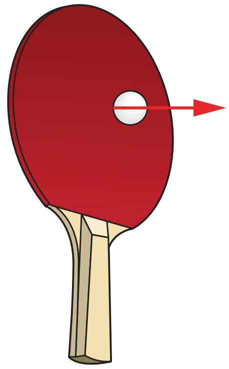
APPLYING NEWTON'S THIRD LAW

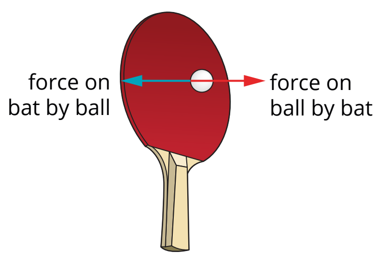
In the diagram below a table-tennis bat is in contact with a table-tennis ball, and one of the forces is given.

a Label the given force using the system ‘F_{on} ________________ by ________________’.

b Label the reaction force to the given force using the system ‘F_{on} ________________ by ________________’.

c Draw the reaction force on the diagram, showing its size and location.



Thinking	**Working**
Identify the two objects involved in the action–reaction pair.	The bat and the ball.
Identify which object is applying the force and which object is experiencing the force, for the force vector shown.	The force vector shown is a force from the bat on the ball.
Use the system of labelling action and reaction forces ‘F_{on} ________________ by ________________’ to label the action force.	$F_{\text{on ball by bat}}$
Use the system of labelling action and reaction forces ‘F_{on} ________________ by ________________’ to label the reaction force.	$F_{\text{on bat by ball}}$
Use a ruler to measure the length of the action force and construct a vector arrow in the opposite direction with its tail on the point of application of the reaction force.	

Worked example: Try yourself 8.5.1

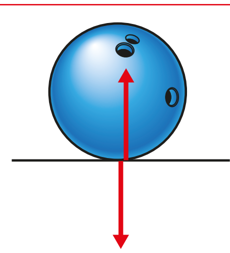
APPLYING NEWTON'S THIRD LAW

In the diagram below, a bowling ball is resting on the floor, and one of the forces is given. Copy the diagram into your book and complete the following:

a Label the given force using the system ‘F_{on} ________ by ________’.

b Label the reaction force to the given force using the system ‘F_{on} ________ by ________’.

c Draw the reaction force on the diagram, showing its size and location.



NEWTON'S THIRD LAW AND MOTION

Newton's third law also explains how you are able to move around. In fact, Newton's third law is needed to explain all motion. Consider walking. Your leg pushes backwards on the ground with each step. This is an action force on the ground by the foot. As shown in Figure 8.5.4, a component of the force acts downwards and another component pushes backwards horizontally along the surface of the ground.

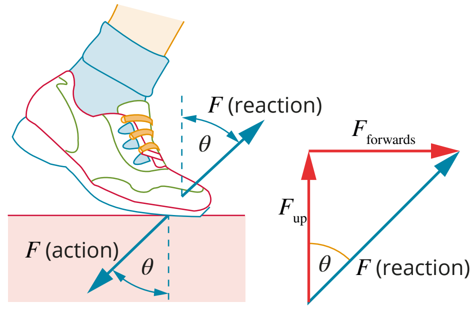


FIGURE 8.5.4 Walking relies on an action and reaction force pair in which the foot pushes down and backwards with an action force. In response, the ground pushes upwards and forwards on you.

The force is transmitted because there is friction between your shoe and the ground. In response, the ground then pushes forwards on you via your foot. This forwards component of the reaction force enables you to move forwards. In other words, it is the ground pushing forwards on you that moves you forwards. It is important to remember that in Newton's second law, $F_{net} = ma$, the net force, F_{net}, is the sum of the forces acting on the body. This does not include forces that are exerted by the body on other objects. When you push back on the ground, this force is acting on the ground and may affect the ground's motion. If the ground is firm, this effect is usually not noticed, but if you run along a sandy beach, the sand is clearly pushed back by your feet.

The act of walking relies on there being some friction between your shoe and the ground. Without it, there is no grip and it is impossible to supply the action force to the ground. Consequently, the ground cannot supply the reaction force needed to enable forwards motion. Walking on smooth ice is a good example of this. Mountaineers use crampons (a rack of sharp spikes) attached to the soles of their boots in order to gain a better grip in icy conditions.

All motion can be explained in terms of action and reaction force pairs. Table 8.5.2 gives some examples of the action and reaction pairs in familiar motions.

TABLE 8.5.2 Action and reaction force pairs are responsible for all types of motion.

Motion	Action force	Reaction force
swimming	hand pushes back on water	water pushes forwards on hand
jumping	legs push down on Earth	Earth pushes up on legs
bicycle or car	tyre pushes back on ground	ground pushes forwards on tyre
jet aircraft and rockets	hot gas is forced backwards out of engine	gases push craft forwards
skydiving	force of gravitation on the skydiver from Earth	force of gravitation on Earth from skydiver

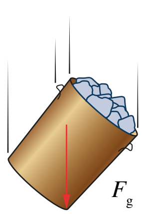


FIGURE 8.5.5 When the bin is in mid-air, there is an unbalanced force due to gravity acting on it so it accelerates towards the ground.

THE NORMAL FORCE

When an object, for example the rubbish bin shown in Figure 8.5.5, is allowed to fall under the influence of gravity, it is easy to see the effect of the force due to gravity. The action force is the force due to gravity of the Earth on the bin, so the net force on the bin is equal to the force due to gravity, and the bin therefore accelerates at $-9.80\,\mathrm{m\,s^{-2}}$.

When the bin is at rest on a table, as shown in Figure 8.5.6a, the force due to gravity ($F_g = mg$) is still acting between the Earth and the bin. Since the bin is at rest, there must be another force acting to balance the force due to gravity on the bin. This upwards force is provided by the table. Because gravity pulls down on the mass of the bin, the bottom of the bin will push down on the surface of the table and the table provides a reaction force on the bin that is equal and opposite, so it will push upwards on the bin as shown in Figure 8.5.6b.

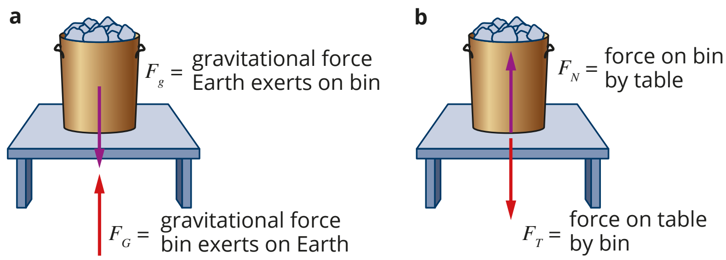


FIGURE 8.5.6 (a) Action–reaction gravitation forces between the bin and the Earth. (b) Action–reaction contact forces between the bin and the table.

The magnitude and direction of the gravitational force on the bin by Earth is equivalent to the magnitude and direction of the force on the table by the bin. Therefore the gravitational force on the bin by Earth is balanced by the upwards contact reaction force on the bin by the table. It is important to note that these two forces are not the pair of forces described in Newton's third law (Figure 8.5.7). This is because the two forces are both acting on the bin and no pair of Newton's third law force pairs acts on the same object. The contact force provided by a surface that is perpendicular to another surface is called the normal reaction force. It is often abbreviated to normal force, and represented by F_N or N.

The reaction force to the force of gravity acting on the bin is in fact the force of gravitation of the bin acting on the Earth. This force, however, is tiny in comparison to the total mass of the Earth and so, from Newton's second law, the acceleration experienced by the Earth is negligible.

When you consider only the forces acting on the bin, you are left with the force due to gravity on the bin by Earth and the normal force on the bin by the table. These two forces come from two separate Newton's third law pairs of forces.

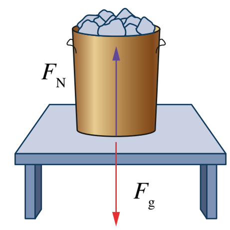


FIGURE 8.5.7 The effect of the two forces on the bin. These are not an action–reaction force pair, even though they are equal in magnitude and opposite in direction.

EXTENSION

The inclined plane

In the example of the bin and the table, the table's surface is horizontal. It is possible that an object could be placed on a surface that is tilted so that it makes an angle, θ, to the horizontal. In this case, the weight force remains the same: $F_g = mg$ downwards. However, the normal force continues to act at right angles to the surface and will change in magnitude, getting smaller as the angle increases. The magnitude of the normal force is equal in size but opposite in direction to the component of the weight force that acts at right angles to the surface. So, the normal force is $F_N = mg \cos \theta$.

The component of the weight force that acts parallel to the surface will cause the mass to slide down the incline. The motion of the object along the plane will be affected by friction, if it is present. The component of the weight force that acts along the surface is given by $F = mg \sin \theta$.

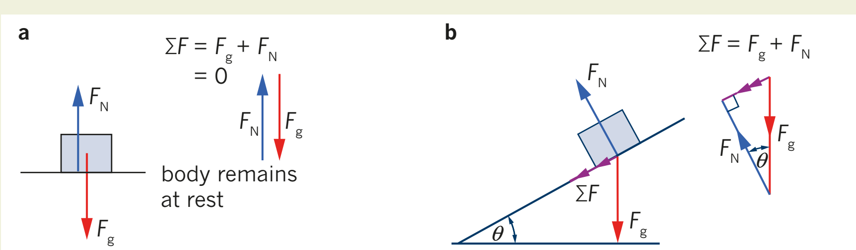


FIGURE 8.5.8 (a) Where the surface is perpendicular to the force due to gravity, the normal force acts directly upwards. (b) On an inclined plane, F_N is at an angle to Fg and is given by $F_N = mg \cos \theta$. If no friction acts, the force that causes the object to accelerate down the plane is $F = mg \sin \theta$.

8.5 Review

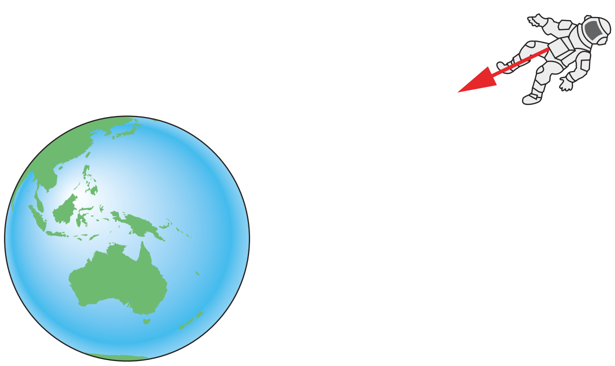
SUMMARY

- For every action (force), there is an equal and opposite reaction (force). This is known as Newton's third law.
- If the action force is labelled systematically, the reaction force can be described by reversing the label of the action force.
- The action and reaction forces are equal and opposite even when the masses of the colliding objects are very different.
- The individual forces making up a Newton's third law pair act on different masses to cause different accelerations according to Newton's second law.
- When an object exerts a downwards force on a surface there is an equal and opposite Newton's third law reaction that exerts a force upwards. This is called the normal force.

KEY QUESTIONS

1 What forces act on a hammer and a nail when a heavy hammer hits a small nail?

2 In the figure below, an astronaut is orbiting the Earth, and one of the forces acting on him is shown by the red arrow.



a Name the given force using the system 'F_{on} ________ $_{\text{by}}$ ________'.

b Name the reaction force using the system 'F_{on} ________ $_{\text{by}}$ ________'.

3 A swimmer completes a training drill in which he doesn't use his legs to kick, but only uses his stroke to move down the pool. What force causes a swimmer to move forwards down the pool?

4 When an inflated balloon is released it will fly around the room. What is the force that causes the balloon to move?

5 Determine the reaction force involved when a ball is hit with a racquet with a force of 100 N west.

6 A 70 kg fisherman is fishing in a 40 kg dinghy at rest on a still lake when, suddenly, he is attacked by a swarm of wasps. To escape, he leaps into the water and exerts a horizontal force of 140 N north on the boat.

a What force does the boat exert on the fisherman?

b With what acceleration will the boat move initially?

c If the force on the fisherman lasted for 0.50 s, determine the initial speed attained by both the man and the boat.

7 An astronaut becomes untethered during a space-walk and drifts away from the spacecraft. To get back to the ship, she decides to throw her tool kit away. In which direction should she throw the tool kit?

8 Two students, James and Tania, are discussing the forces acting on a lunchbox that is sitting on the laboratory bench. James states that the weight force and the normal force are acting on the lunchbox and that since these forces are equal in magnitude but opposite in direction, they comprise a Newton's third law action–reaction pair. Tania disagrees, saying that these forces are not an action–reaction pair. Who is correct and why?

8.6 Impulse and force

Section 8.4 on Newton's second law of motion discussed the quantitative connection between force, mass, time and change in velocity. This relationship is explored further in this section. The relationship between change in momentum, Δp, the period of time, Δt, and net force, F_{net}, helps to explain the effects of collisions and how to minimise those effects. It is the key to providing safer environments, including in sporting contexts such as that shown in Figure 8.6.1.

FIGURE 8.6.1 When two footballers collide, they exert an equal and opposite force on each other.

Think about what it would feel like to fall onto a concrete floor. Even from a small height it would hurt. A fall from the same height onto a foam mattress would barely be felt. In both situations speed is the same, mass has not changed and gravity provides the same acceleration, no matter the mass. Yet each experience would feel different.

CHANGE IN MOMENTUM (IMPULSE)

According to Newton's second law, a net force will cause a mass to accelerate. A larger net force will create a faster change in the velocity of the mass. The faster the change occurs—that is, the smaller the period of time, Δt—the greater the net force that produced that change. Landing on a concrete floor changes the velocity of an object very quickly. The falling object is brought to an abrupt stop within a very short amount of time. When landing on a foam mattress, the change occurs over a much greater timeframe. Therefore, the force needed to produce the change is smaller.

Starting with the equation introduced in Section 8.4, the relationship between change in momentum, Δp, or impulse, I, and the variables of force, F_{net} (often written just as F), and period of time, Δt, becomes:

$$\begin{aligned} F_{net} &= \frac{\Delta p}{\Delta t} \\ &= \frac{mv - mu}{\Delta t} \\ &= \frac{m(v-u)}{\Delta t} \\ F_{net}\Delta t &= m(v-u) \\ &= I \end{aligned}$$

where I is the impulse ($kg\,m\,s^{-1}$).

These equations illustrate that for a given change in momentum or impulse, the product of force and period of time is constant. This relationship is key to understanding collisions. Worked examples 8.6.1 and 8.6.2, below, illustrate how this works.

Worked example 8.6.1

CALCULATING THE FORCE AND IMPULSE

A student drops a 105 g pool ball onto a concrete floor from a height of 2.00 m. Just before it hits the floor, the velocity of the ball is 6.26 m s^{-1} down. Before it bounces back up, there is an instant in time at which the ball's velocity is zero. The time it takes for the ball to change its velocity to zero is 5.02 milliseconds.

a Calculate the change in momentum of the pool ball.

Thinking	Working
Ensure that the variables are in their standard units.	$m = 0.105$ kg $u = 6.26$ m s^{-1} down $v = 0$ m s^{-1}
Apply the sign and direction convention for motion in one dimension. Up is positive and down is negative.	$m = 0.105$ kg $u = -6.26$ m s^{-1} $v = 0$ m s^{-1}
Apply the equation for change in momentum.	$\Delta p = m(v - u)$ $= 0.105 \times (0 - (-6.26))$ $= 0.657$ kg m s^{-1}
Refer to the sign and direction convention to determine the direction of the change in momentum.	$\Delta p = 0.657$ kg m s^{-1} up

b Calculate the impulse of the pool ball.

Thinking	Working
Using the answer to part (a), apply the equation for impulse.	$I = \Delta p$ $= 0.657$ kg m s^{-1}
Refer to the sign and direction convention to determine the direction of the impulse.	$I = 0.657$ kg m s^{-1} up

c Calculate the average force that acts to cause the impulse.

Thinking	Working
Use the answer to part (b). Ensure that the variables are in their standard units.	$I = 0.657$ kg m s^{-1} $\Delta t = 5.02 \times 10^{-3}$ s
Apply the equation for force.	$F\Delta t = I$ $F = \dfrac{I}{\Delta t}$ $= \dfrac{0.657}{5.02 \times 10^{-3}}$ $= 131$ N
Refer to the sign and direction convention to determine the direction of the force.	$F = 131$ N up

Worked example: Try yourself 8.6.1

CALCULATING THE FORCE AND IMPULSE

A student drops a 56.0 g egg onto a table from a height of 60 cm. Just before it hits the table, the velocity of the egg is 3.43 m s^{-1} down. The egg's final velocity is zero as it smashes on the table. The time it takes for the egg to change its velocity to zero is 3.55 ms.

a Calculate the change in momentum of the egg.

b Calculate the impulse of the egg.

c Calculate the average force that acts to cause the impulse.

Worked example 8.6.2

CALCULATING THE FORCE AND IMPULSE (SOFT LANDING)

A student drops a 105 g pool ball onto a foam mattress from a height of 2.00 m. Just before it hits the foam mattress, the velocity of the ball is 6.26 m s^{-1} down. Before it bounces back up, there is an instant in time at which the ball's velocity is zero. The time it takes for the ball to change its velocity to zero is 0.360 seconds.

a Calculate the change in momentum of the pool ball.	
Thinking	**Working**
Ensure that the variables are in their standard units.	$m = 0.105$ kg $u = 6.26$ m s^{-1} down $v = 0$ m s^{-1}
Apply the sign and direction convention for motion in one dimension. Up is positive and down is negative.	$m = 0.105$ kg $u = -6.26$ m s^{-1} $v = 0$ m s^{-1}
Apply the equation for change in momentum.	$\Delta p = m(v - u)$ $= 0.105 \times (0 - (-6.26))$ $= 0.657$ kg m s^{-1}
Refer to the sign and direction convention to determine the direction of the change in momentum.	$\Delta p = 0.657$ kg m s^{-1} up

b Calculate the impulse of the pool ball.	
Thinking	**Working**
Using the answer to part (a), apply the equation for impulse.	$I = \Delta p$ $= 0.657$ kg m s^{-1}
Refer to the sign and direction convention to determine the direction of the impulse.	$I = 0.657$ kg m s^{-1} up

c Calculate the average force that acts to cause the impulse.	
Thinking	**Working**
Using the answer to part (b), ensure that the variables are in their standard units.	$I = 0.657\ \text{kg m s}^{-1}$ $\Delta t = 0.360\ \text{s}$
Apply the equation for force.	$F\Delta t = I$ $F = \dfrac{I}{\Delta t}$ $= \dfrac{0.657}{0.360}$ $= 1.83\ \text{N}$
Refer to the sign and direction convention to determine the direction of the force.	$F = 1.83\ \text{N}$ up

Worked example: Try yourself 8.6.2

CALCULATING THE FORCE AND IMPULSE (SOFT LANDING)

A student drops a 56.0 g egg into a mound of flour from a height of 60 cm. Just before it hits the mound of flour, the velocity of the egg is $3.43\ \text{m s}^{-1}$ down. The egg's final velocwity is zero as it sinks into the mound of flour. The time it takes for the egg to change its velocity to zero is 0.325 seconds.

a Calculate the change in momentum of the egg.

b Calculate the impulse of the egg.

c Calculate the average force that acts to cause the impulse.

From these worked examples you should notice a number of important things:

- The change in momentum and the impulse were always the same.
- Regardless of the surface that the object landed on, the impulse or change in momentum remained the same.
- The period of time was the main difference between the two different surfaces. Hard surfaces resulted in a short time to stop and soft surfaces resulted in a longer time to stop.
- The effect of the period of time on the force was dramatic. A shorter time meant a greater force, while a longer time meant a much smaller force.

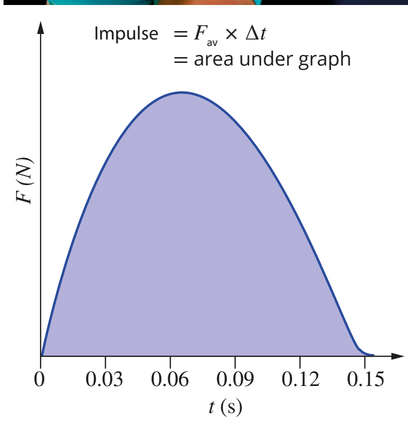


FIGURE 8.6.2 The forces acting on the tennis ball during its collision with the racquet are not constant.

DETERMINING IMPULSE FROM A CHANGING FORCE

In the previous examples it was assumed that the force that acted to change the impulse over a period of time was constant during that time. This is not always the case in real situations. Often the force varies over the period of the impact, so there needs to be a way to determine the impulse as the force varies.

An illustration of this is when a tennis player strikes a ball with a racquet. At the instant the ball comes in contact with the racquet, the applied force will be small. As the strings distort and the ball compresses, the force will increase until the ball has been stopped. The force will then decrease as the ball accelerates away from the racquet. A graph of force against time is shown in Figure 8.6.2.

The impulse, I, affecting the ball during any time interval will be the product of applied force, F, and the period of time, Δt. The total impulse during the period of time the ball is in contact with the racquet will be:

$$I = F_{av}\Delta t$$

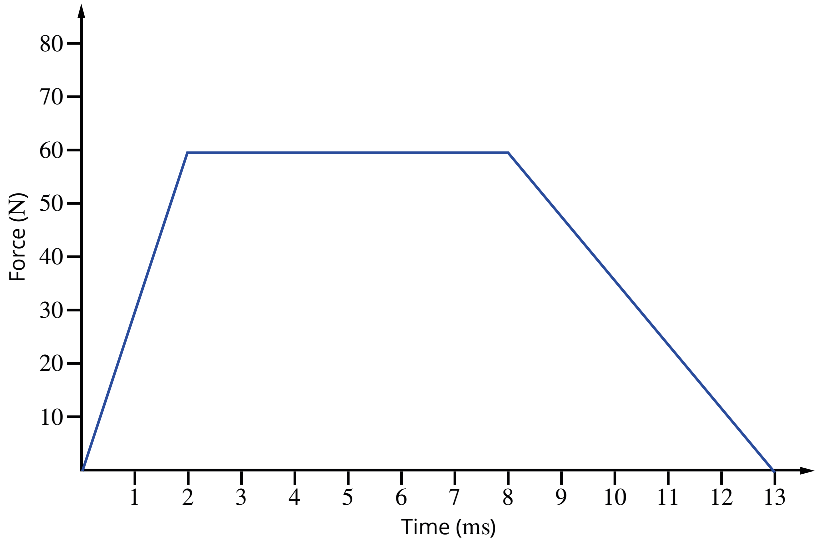
where F_{av} is the average force applied during the collision and Δt is the total period of time the ball is in contact with the racquet. In a graph showing force against time, the area under the line is some product of the height (force) and the width (time). Thus, the total area under the line in a force against time graph is the total impulse for any collision, even those in which the force is not constant.

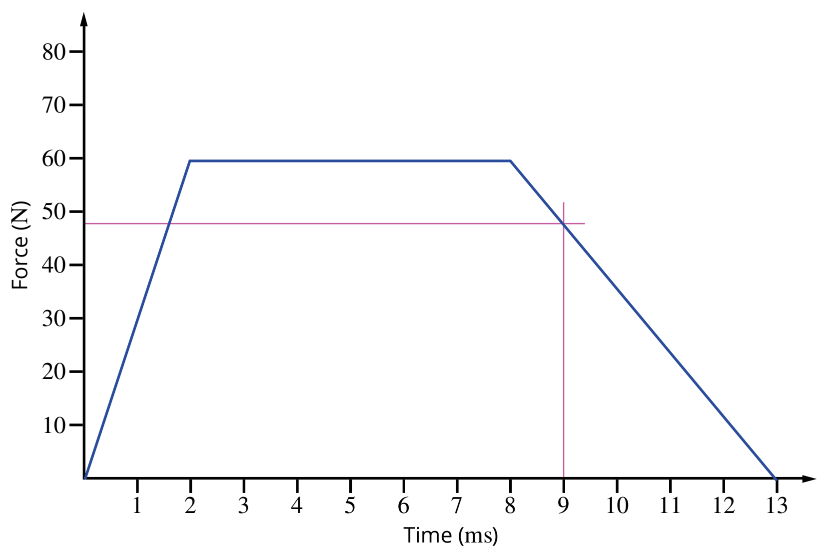
The concept of impulse is appropriate when dealing with forces during any collision since it links force and contact time as, for example, when a person's foot hits the ground or when a ball is hit by a bat or racquet. If applied to situations where contact is over an extended period of time, the average net force involved is used since the forces are generally changing (as the ball deforms, for example). The average net applied force can be found directly from the formula for impulse. The instantaneous applied force at any particular time during the collision must be read from a graph of force against time.

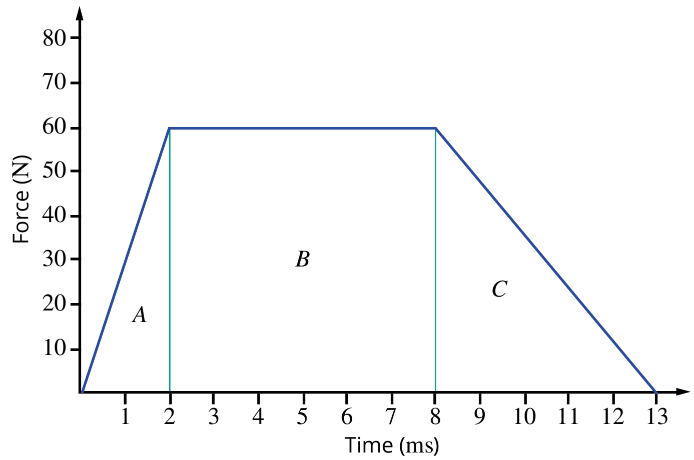
Worked example 8.6.3

CALCULATING THE TOTAL IMPULSE FROM A CHANGING FORCE

A student records the force acting on a rubber ball as it bounces off a hard concrete floor over a period of time. The graph shows the forces acting on the ball during its collision with the concrete floor.



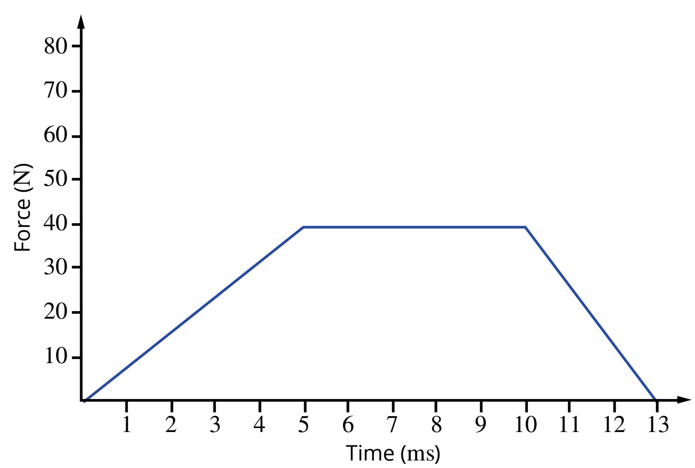
a Determine the force acting on the ball at a time of 9.0 ms.	
Thinking	**Working**
From the 9.0 ms point on the *x*-axis go up to the line of the graph, then across to the *y*-axis.	
The force is estimated by reading the intercept of the *y*-axis.	$F = 48\,\text{N}$

b Calculate the total impulse of the ball over the 13 ms period of time.	
Thinking	**Working**
Break the area under the graph into sections for which you can calculate the area.	In this case, the graph can be broken into three sections: *A*, *B* and *C*. 
Calculate the area of the three sections *A*, *B* and *C* using the equations for area of a triangle and the area of a rectangle.	area = $A + B + C$ $= \left(\frac{1}{2} b \times h\right) + (b \times h) + \left(\frac{1}{2} b \times h\right)$ $= \left[\frac{1}{2} \times (2.0 \times 10^{-3}) \times 60\right] + \left[(6.0 \times 10^{-3}) \times 60\right] + \left[\frac{1}{2} \times (5.0 \times 10^{-3}) \times 60\right]$ $= 0.060 + 0.36 + 0.15$ $= 0.57$
The total impulse is equal to the area.	I = area $= 0.57\ \text{kg m s}^{-1}$
Apply the sign and direction convention for motion in one dimension vertically.	$I = 0.57\ \text{kg m s}^{-1}$ up

Worked example: Try yourself 8.6.3

CALCULATING THE TOTAL IMPULSE FROM A CHANGING FORCE

A student records the force acting on a tennis ball as it bounces off a hard concrete floor over a period of time. The graph shows the forces acting on a ball during its collision with the concrete floor.



a Determine the force acting on the ball at a time of 4.0 ms.

b Calculate the total impulse of the ball over the 13 ms period of time.

PHYSICS IN ACTION

Car safety

Designing a successful car is a complex task. A vehicle must be reliable, economical, powerful, visually appealing, secure and safe. Public perception of the relative importance of these issues varies. Magazines and newspapers concentrate on appearance, price and performance. The introduction of air-bag technology into most cars has altered the focus towards safety.

Vehicle safety is primarily about crash avoidance. Research shows potential accidents are avoided 99% of the time. The avoidance of accidents is mainly due to accident avoidance systems such as antilock brakes. When a collision does happen, passive safety features, such as the air bag, come into operation. Understanding the theory behind accidents involves primarily an understanding of impulse and force.

The seatbelt

Whenever you enter a car, your first instinct is probably to reach for the seatbelt. Seatbelts started being included in cars in the 1950s, and in 1970 Victoria introduced the world's first seatbelt law.

The design of seatbelts works directly to try to mitigate the effects of Newton's first law. In order to get from one place to another at high speeds in a car, the occupants of the vehicle must themselves be travelling at high speeds within the car. You have seen from Newton's first law that a body in motion stays in motion unless acted on by an unbalanced force. This means that if a car is involved in a collision, the occupants of the vehicle will keep moving, resulting in either ejection from the car or contact with the interior—causing serious injury or death. The seatbelt is designed to lock during severe deceleration and acts to oppose the occupants' motion. Seatbelts are a vital component of car safety—experts estimate that they decrease the risk of fatality by up to 50%.

The airbag

The introduction of seatbelts allowed many more people to survive car accidents. However, many of these survivors sustained serious injuries. So, although seatbelts saved lives, there was also an increase in serious injuries. A further safety device was required to minimise these injuries.

The airbag in a car is designed to inflate within a few milliseconds of the occurrence of a collision to reduce secondary injuries during the collision. The airbag is designed to inflate only when the vehicle experiences an impact with a solid object at 18–20 km h^{-1} or more. The required deceleration must be high, or accidental nudges with another car would cause the airbag to inflate. The car's computer control makes a decision within a few milliseconds to detonate the gas cylinders that inflate the airbag. The propellant detonates and inflates the airbag. According to Newton's first law, the driver continues to move towards the dashboard. As the driver continues forwards into the airbag, the bag deflates, allowing the body to slow down over a longer time than would otherwise be possible as it moves towards the dashboard (Figure 8.6.3). The force is minimised so injury is reduced.

FIGURE 8.6.3 Airbags can prevent injuries by extending the period of time you take to stop.

Car safety *continued*

Calculating exactly when the airbag should inflate, and for how long, is a difficult task. Many cars have been crash tested and the results painstakingly analysed. High-speed film demonstrates precisely why the airbag is so effective. During a collision the arms, legs and head of the occupants are restrained only by the joints and muscles. Enormous forces are involved because of the large deceleration. Looking back to the concept of impulse, the driver's net change in momentum will be exactly the same as if they came to a halt themselves. However, the time over which the impulse takes place is reduced to milliseconds. From $I = F\Delta t$ it can be seen that this can increase the total force to a huge amount. The shoulders and hips can, in most cases, sustain the large forces for the short duration. However, the neck is the weak link. Victims of road accidents regularly receive neck and spinal injuries. An airbag reduces the enormous forces the neck must withstand by extending the duration of the collision. This involves the direct application of the concept of impulse. A comparison of the forces applied to the occupant of a car with and without airbags is shown in Figure 8.6.4.

Airbags prevent the high forces caused by contact of the head with the steering wheel. The airbag ensures that the main thrust of the expansion is directed outwards instead of towards the driver. The airbag's deflation rate, governed by the size of the holes in the rear of the air bag, provides the optimum deceleration of the head for a large range of impact speeds.

The airbag is not the answer to all safety concerns associated with a collision, but it is one of many safety features that form a chain of defence in a collision.

Crumple zones, helmets and safety barriers

Crumple zones, helmets and safety barriers work on similar principles to airbags. Humans are susceptible to damage from large forces, and so in the event of an unavoidable accident, the primary goal of safety features is to reduce the amount of force acting on the person by extending the duration of the collision.

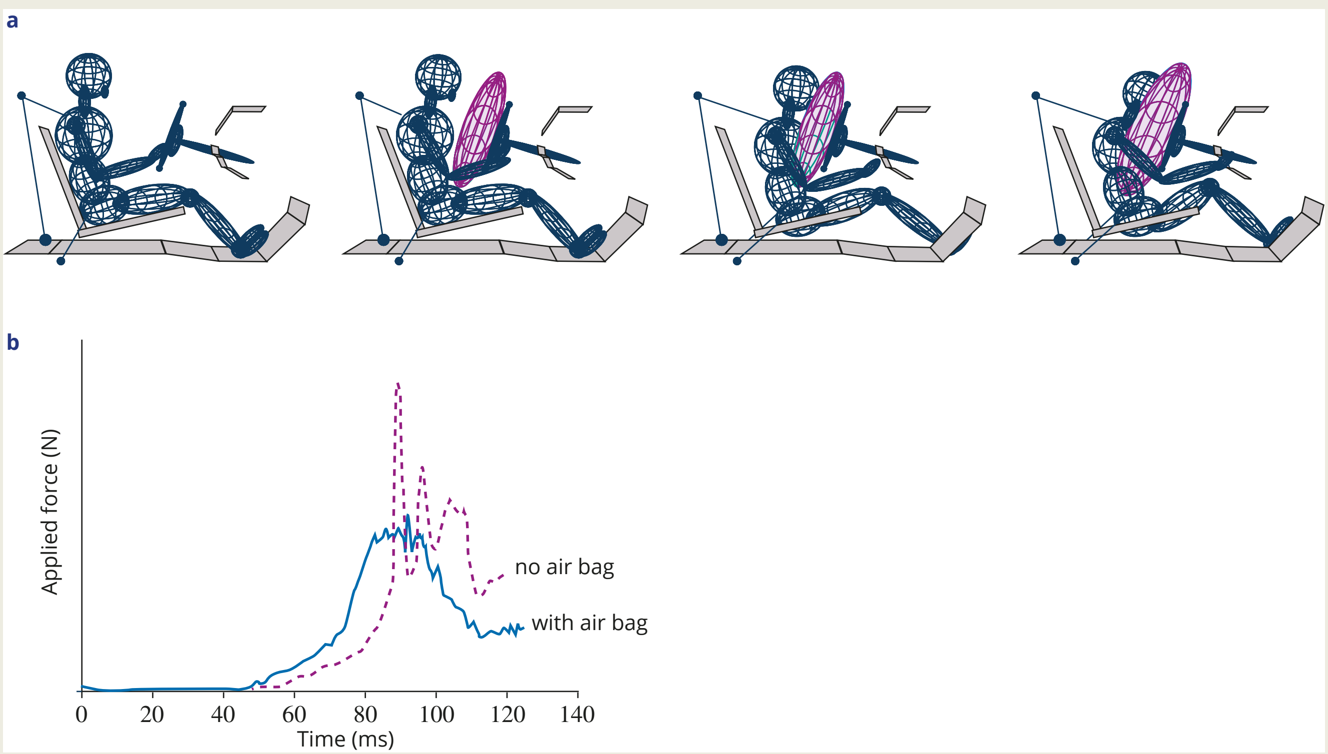


FIGURE 8.6.4 (a) The airbag extends the stopping time and distributes the force required to decelerate the mass of the driver or passenger over a larger area than a seatbelt. (b) The force withstood by the occupant of the car without an airbag is about double that felt with an airbag.

Crumple zones are a structural feature in most cars (Figure 8.6.5). By providing a large distance between the driver and the front of the vehicle in the form of a long bonnet made out of malleable metals, engineers can extend the time period over which a crash takes place. Therefore, the total force applied is lowered before the impact has even reached the driver. Furthermore, much of the deformation which occurs to the car allows the energy of the crash to be dissipated into the metal rather than being delivered directly to the driver and passengers.

FIGURE 8.6.5 The crumple zone on this car has prevented serious damage to the car's occupants.

Helmets are similar in principle to crumple zones. The outer layer is typically made with a type of crushable foam. It controls the energy of an impact into the helmet itself rather than the user's head, extending collision time and spreading the energy of the collision over a greater area. The second statement, whereby the impact is spread out over the head rather than being restricted to one area, is perhaps a little more intuitive to you than the first, which will require some explanation. Conservation of energy tells us that in a collision, all of the kinetic energy present has to either remain present, or be transformed into another kind of energy. In the absence of any protective gear, this means that the very high kinetic energy of the collision will all be transferred into a very localised part of a rider's head. Helmets allow some of this kinetic energy to be spread out safely into the helmet to reduce the energy that reaches the head. Furthermore, helmets are usually designed with soft, padded interiors. This, in addition to the compressible foam, allows the helmet to move around a little in the event of an impact, therefore increasing the time of collision in the same way as airbags and crumple zones, and reducing the total force to which the rider is exposed.

You may have seen traffic safety barriers at the side of roads and wondered what their purpose is—after all, a crash into a safety barrier will typically be anything but safe. They are designed to stop out-of-control vehicles running off the road or travelling into the path of other traffic (Figure 8.6.6). You know from Newton's first and second laws that a moving body needs an opposing force in order to stop it, and that in order to stop something moving very fast in a short amount of time the opposing force needs to be very large. Safety barriers are designed to withstand a high amount of force—this allows them to provide the normal force required back at vehicles in order to stop their motion. The sudden impact can be very damaging to the out-of-control vehicle, but ultimately the purpose of safety barriers is to prevent more widespread harm.

FIGURE 8.6.6 A steel safety barrier protecting motorists on a freeway bridge.

8.6 Review

SUMMARY

- Newton's second law describes the relationship between impulse, force and the period of time: $I = F\Delta t$
- The same mass changing its velocity by the same amount will have a constant change in momentum or impulse.
- The faster a mass changes its velocity, the greater the force required to change the velocity in that period of time.
- The slower a mass changes its velocity, the smaller the force required to change the velocity in that period of time.
- Forces can change during a collision.
- The impulse over a period of time can be found by calculating the area under the line on a force versus time graph.
- The period of time is the cause of the difference between the force provided by two different surfaces during a collision. Hard surfaces result in a short time to stop and soft surfaces result in a longer time to stop.
- The effect of the period of time on the force is dramatic. A shorter time means a greater force, while a longer time means a much smaller force.

KEY QUESTIONS

1 A 45.0 kg mass changes its velocity from 2.45 m s^{-1} east to 12.5 m s^{-1} east in a period of 3.50 s.
- **a** Calculate the change in momentum of the mass.
- **b** Calculate the impulse of the mass.
- **c** Calculate the force that causes the impulse of the mass.

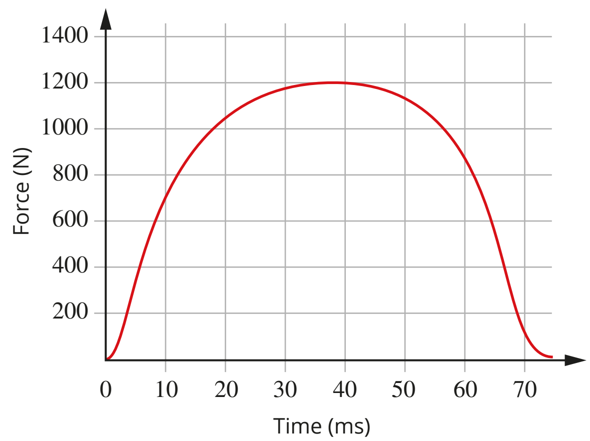
2 Using the concept of impulse, explain how airbags can reduce injuries during a collision.

3 A student catches a 75.0 g cricket ball with 'hard hands'. Just before the student catches the ball, its velocity is 15.6 m s^{-1} west. With hard hands, the velocity of the ball drops to zero in just 0.100 seconds.
- **a** Calculate the change in momentum of the cricket ball.
- **b** Calculate the impulse of the cricket ball.
- **c** Calculate the average force on the cricket ball.

4 The student from question 3 now catches the same 75.0 g cricket ball, but this time with 'soft hands'. Just before the student catches the ball, its velocity is 15.6 m s^{-1} west. With soft hands, the velocity of the ball drops to zero in 0.300 seconds. Calculate the average force on the cricket ball.

5 A 200 g cricket ball (at rest) is struck by a cricket bat. The ball and bat are in contact for 0.05 s, during which time the ball is accelerated to a speed of 45 m s^{-1}.
- **a** What is the magnitude of the impulse the ball experiences?
- **b** What is the net average force acting on the ball during the contact time?
- **c** What is the net average force acting on the bat during the contact time?

6 The following graph shows the net vertical force generated as an athlete's foot strikes an asphalt running track.



- **a** Estimate the maximum force acting on the athlete's foot during the contact time.
- **b** Estimate the total impulse during the contact time.

7 A 25 g arrow buries its head 2 cm into a target on striking it. The arrow was travelling at 50 m s^{-1} just before impact.
- **a** What change in momentum does the arrow experience as it comes to rest?
- **b** What is the impulse experienced by the arrow?
- **c** What is the average force that acts on the arrow during the period of deceleration after it hits the target?

8 Crash helmets are designed to reduce the force of impact on the head during a collision.
- **a** Explain how their design reduces the net force on the head.
- **b** Would a rigid 'shell' be as successful? Explain.

8.7 Mass and weight

The difference between mass and weight is sometimes misunderstood because the terms are used interchangeably in the English language. In physics, however, the two terms have different meanings. For instance, weight is a vector and mass is a scalar. The difference between these two terms is explained in this section.

MASS OF A BODY

Mass is a scalar quantity. In scientific contexts, mass is measured in kilograms (kg). In earlier science courses, **mass** may have been defined as 'the amount of matter in an object'. Since the late 1700s, the kilogram has been defined in terms of an amount of a standard material. At first, 1 litre of water at 4°C was used to define the kilogram. More recently an international mass standard has been introduced. This is a 1 kg cylinder of platinum–iridium alloy that is kept in Paris (Figure 8.7.1). Copies are made from this standard and sent around the world to calibrate balances.

FIGURE 8.7.1 All mass in the world is compared to this small piece of platinum–iridium alloy held in a sealed vault in Paris.

PHYSICSFILE

Defining a kilogram

The kilogram is currently defined by the mass of a cylinder made of platinum–iridium alloy kept in Paris. It is the only SI unit that is defined in this way. A group of scientists from around the world, including Australian scientists, are currently looking at ways to redefine the kilogram in terms of a physical property that is unchanging and can be reproduced in laboratories. The Avogadro Project has been running for several years with the aim that soon there will be a scientific definition of the kilogram.

FIGURE 8.7.2 CSIRO scientists are working on a replacement for the standard kilogram.

You have already encountered an understanding of mass in Section 8.4 with the introduction of Newton's second law and the effect of a force on a massive body. An object's mass, m, can be seen as a property by the amount of acceleration it undergoes for a given net force F. The more massive an object is, the less it will accelerate for that given force. Therefore, the concept of mass is directly tied with the *inertia* of that object. That is, the mass of an object can be seen as its unwillingness to move when a force is applied.

The more mass an object has, the greater the force required to make it accelerate. If the same force is applied to two different masses, the smaller mass will accelerate more than the greater mass. For this reason, mass can be seen as the property of a body that resists the change in motion caused by a force.

If the above experiment is repeated on the Moon with the same horizontal force acting on the body on a frictionless surface, the same acceleration will result, as shown in Figure 8.7.3. This is because the mass of the body remains the same on the Earth and on the Moon. Mass is a property of the body and it is not affected by its environment.

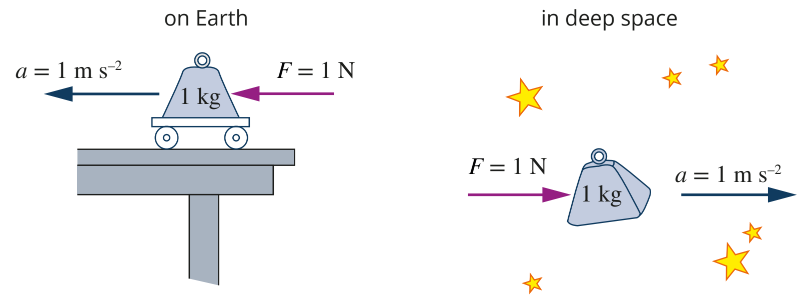


FIGURE 8.7.3 Mass is a property of the object and not of its surroundings.

GRAVITATIONAL FORCE

In the late 1500s, Galileo was able to show that all objects which are dropped near the surface of the Earth accelerate at the same rate, g, towards the centre of the Earth. The force that produces this acceleration is the force due to gravity. The force due to gravity is an attractive force that exists between all masses. In other words, it is a 'pulling' force that exists between everything that has a mass. It is one of the fundamental forces that acts over a distance, which means that the two masses do not need to be in contact in order for the force to exist.

Gravitational forces result from a mass creating a gravitational field that spreads throughout the space around the mass. Any other mass that is within this field will experience a force towards the mass creating the field. The object that is in the gravitational field also has mass, so it too has a gravitational field around it that attracts the original mass with an equal and opposite force.

The gravitational field extends through space in an inverse squared relationship. This means that if you double the distance from the mass creating the field, then the force will be one-quarter the size.

Here on Earth, you are strongly affected by the Earth's gravitational field (and less strongly by fields from the Sun, the Moon and other objects in the solar system). Even if you were not on Earth, you could still measure the effect of the Earth's gravitational field. There is no place in the universe where the Earth's gravitational field will not reach. At the 'edge' of the universe it will be very small, but it can be calculated. The closer you or any mass is located to the Earth, the larger the gravitational force of attraction towards the Earth. At a height above the Earth's surface that is equal to the radius of the Earth, the force due to gravity on a mass will be one-quarter of that at the Earth's surface. At two Earth radii above the Earth's surface, the gravitational force will be one-ninth of the force experienced on the Earth's surface.

PHYSICSFILE

Inertial mass and gravitational mass

You may be wondering why there are two different definitions or indications of the mass of an object: its resistance to movement and the force exerted on it by gravity. In fact, these are two conceptually distinct properties of the object; there is no known fundamental or compelling reason why the gravitational mass should be exactly equal to inertial mass, it just seems to be that way.

To a precision of 10^{-12} kg researchers have been unable to find a difference in value between the two different types of mass. While the notion is interesting, you don't need to worry about these differences in any way. Because the values are the same to a very high degree of precision, it is completely acceptable to treat them as the exact same quantity. However, finding a consistent underlying reason to their relationship remains an active topic of research in physics today.

PHYSICSFILE

The force of gravity between the Earth and the Moon

The Moon stays in orbit around the Earth due to the attractive force of gravity that acts between the two bodies. The force required to maintain the Moon in an orbit of the Earth is very large. If it could be replaced with a steel rod that connects the Earth and the Moon, the rod would have to be over 700 km in diameter.

WEIGHT FORCE

In physics, the force on a body due to gravity is called the **weight** of a body, F_g or just W. Weight is a force, therefore it is a vector quantity. Like other forces, it is measured in newtons (N).

Figure 8.7.4 shows a bin falling through the air. As it falls, it accelerates vertically downwards due to the Earth's gravitational field strength, g, which near the surface of the Earth is $9.80\,\mathrm{N\,kg^{-1}}$ down.

As the weight of the bin is a vector, it can be represented with an arrow. An arrow representing weight is drawn downwards (towards the centre of the Earth) with its tail beginning at the object's centre of mass. Centre of mass is described in detail in chapter 7. Stated simply, an object's centre of mass is the point where its mass can be considered to be 'concentrated'. In an object of uniform density, there is as much mass above the centre of mass as there is below it, as much mass to the left as there is to the right, and as much mass in front as there is behind it.

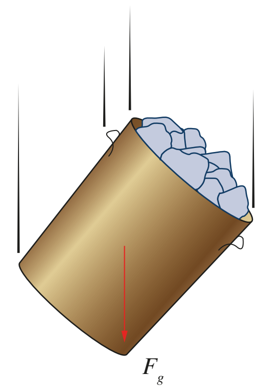


FIGURE 8.7.4 When the bin is in mid-air, there is an unbalanced force due to gravity acting on it so it accelerates towards the ground. This is the object's weight force. The vector representing weight is drawn from the centre of mass of the object and points downwards.

> The *weight* of a body F_g (in N) is defined as the force of attraction on a body due to gravity and is calculated using the equation:
>
> $F_g = mg$
>
> where F_g is the force of gravity acting at the centre of mass of a body (in N)
>
> m is the mass of the body (in kg)
>
> g is the gravitational field strength (in $\mathrm{N\,kg^{-1}}$, which is $9.8\,\mathrm{N\,kg^{-1}}$ near the surface of the Earth).

Your mass and weight on the Moon

If you were ever lucky enough to travel to the Moon, you would notice that over the duration of the trip, the amount of matter that makes up your body would not change. There would still be as much of you present when you arrived on the Moon as there was when you left the Earth. Your mass wouldn't have changed because mass is a property of the matter and is not affected by its environment. However, you would notice when you stood on the floor of the Moon base that you were not pulled down as hard on the ground. In other words, your weight force would be much less on the Moon than it was on the Earth.

The Moon has a much smaller mass than the Earth. The Moon's mass is $7.35 \times 10^{22}\,\mathrm{kg}$ and the Earth's mass is $5.97 \times 10^{24}\,\mathrm{kg}$. This means that the Moon's mass is about 81 times smaller than the Earth. This smaller mass, and the radius of the Moon, means that the Moon's gravitational field is much weaker than the Earth's gravitational field and therefore the force due to gravity on any object is much less. As the force is less, the acceleration that an object experiences on the Moon will be less than here on Earth. Consider Figure 8.7.5, in which a 5 kg pumpkin is falling on the Earth and then on the Moon.

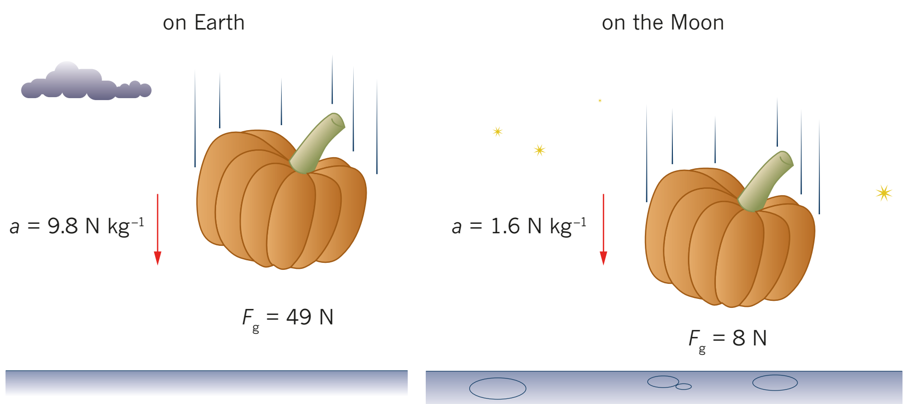


FIGURE 8.7.5 A 5 kg pumpkin falling on the Earth and on the Moon. The mass of the pumpkin is 5 kg no matter where it is, but the weight of the pumpkin is different.

The pumpkin has a much smaller weight force on the Moon than on the Earth due to the much smaller gravitational field.

8.7 Review

SUMMARY

- The standard mass is a 1 kg platinum–iridium cylinder, against which all other masses are compared.
- The mass of an object relates to its ability to resist changes in motion.
- Mass is a scalar quantity and is a property of a body that is not influenced by external environmental factors.
- Gravity is a force of attraction between masses that extends throughout space.
- Weight is a force due to gravity. As it is a force, it is also a vector and requires a magnitude and direction.
- Mass is measured in kilograms and weight is measured in newtons.
- The mass of an object will not change as the object goes from the Earth to the Moon, but the weight will change due to the decreased gravitational force on the Moon compared to the Earth.
- The centre of mass indicates the position at which the entire mass of a body is considered to be concentrated. At this point, all external forces are applied.

KEY QUESTIONS

1 An object is placed in a spaceship and launched into space. As the object leaves the surface of the Earth, it has a mass of 50 kg. What will be the object's mass if it lands on Pluto?

2 Two students have a conversation. One of the students states that her weight is 60 kg. What is wrong with this statement?

3 Mary's mass is 75 kg. What is her weight on Earth if g is $9.8\ \text{N kg}^{-1}$?

4 A desk chair has a weight of 34.3 N on the surface of the Earth. Determine the mass of the chair. Use $g = 9.8\ \text{N kg}^{-1}$.

5 Determine the weight of the same chair in the previous question when it is on the surface of the Moon, where $g = 1.6\ \text{N kg}^{-1}$.

6 On the surface of the Earth, a geological hammer has a mass of 1.5 kg. Determine its mass and weight on Mars, where $g = 3.6\ \text{N kg}^{-1}$.

7 Would your weight be greater on Earth or on the Moon? Explain your answer.

Chapter review

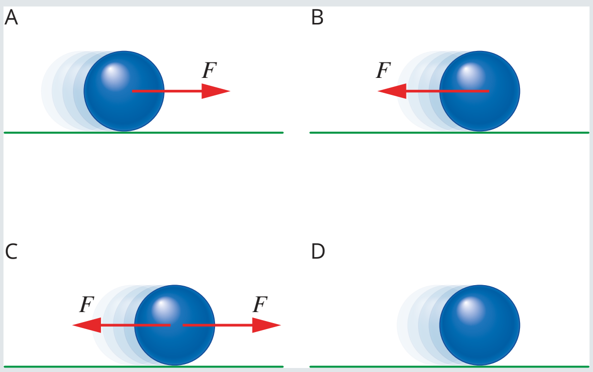
08

KEY TERMS

conserved
contact forces
force
impulse
inertia
mass
momentum
net force
newton
Newton's first law
Newton's second law
Newton's third law
non-contact forces
weight

1 A student is travelling to school on a train. When the train starts moving, she notices that passengers tend to lurch towards the back of the train before grabbing a handrail to stop themselves from falling. Has a force acted to push the passengers backwards? Justify your answer.

2 A bowling ball rolls along a smooth wooden floor at constant velocity. Ignoring the effects of friction and air resistance, which of the following diagrams correctly indicates the forces acting on the ball?



3 Calculate the mass of an object if it accelerates at 9.20 m s^{-2} east when a force of 352 N east acts on it.

4 Calculate the acceleration of a 657 kg motorbike when a net force of 3550 N north acts on it.

The following information relates to questions 5–8.

Lachy is riding his bike and producing a forwards force of 150 N. The combined mass of Lachy and the bike is 100 kg.

5 If there is no friction or air resistance, what is the magnitude of the acceleration of Lachy and the bike?

6 If friction opposes the bike's motion with a force of 45.0 N, what is the magnitude of the acceleration of the bike?

7 What must be the magnitude of the force of friction if Lachy's acceleration is 0.600 m s^{-2}?

8 Lachy now carries an additional mass of 25.0 kg due to his school bag. What must be the new forwards force he produces in order to accelerate at 0.800 m s^{-2} if friction opposes the motion with a force of 30.0 N?

9 Calculate the final velocity of a 14.0 kg remote-controlled car moving at 3.75 m s^{-1} east, when a net force of 62.0 N west acts on it for 2.00 seconds.

10 A 65.0 kg student is standing on a 3.50 kg skateboard at rest when he steps off the board and exerts a horizontal force of 75.0 N south on the board. What force does the board exert on the student?

11 Calculate the change in momentum of a 155 kg rock whose velocity has changed from 6.50 m s^{-1} east to 3.25 m s^{-1} east in a period of 8.50 s.

12 Calculate the change in momentum of a 25.5 kg robot whose velocity changes from 6.40 m s^{-1} forwards to 2.25 m s^{-1} backwards.

13 An astronaut in a protective suit has a total mass of 154 kg and throws a 40.0 kg toolbox away from the space station. The astronaut and toolbox are initially stationary. After being thrown, the toolbox moves at 2.15 m s^{-1}. Calculate the velocity of the astronaut just after throwing the toolbox.

14 A 75.0 kg netball player is moving at 4.00 m s^{-1} west changes direction during a game to 5.00 m s^{-1} north. Calculate the change in momentum of the player over the period of the change.

15 An athlete catches a 300 g volleyball by relaxing her elbows and wrists and 'giving' with the ball. Just before the athlete catches the ball, its velocity is 5.60 m s^{-1} west. With soft hands, the velocity of the ball drops to zero in 1.00 s. Calculate the average force exerted by the athlete on the volleyball.

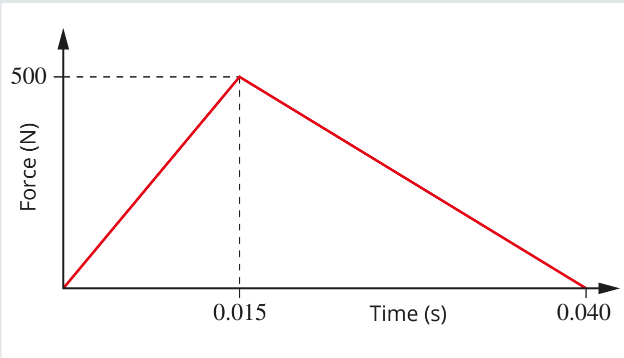
16 Explain how crumple zones reduce the damage done to occupants of a car in the event of a collision. Is it better to have a completely rigid metal bonnet, a completely soft metal bonnet, or somewhere in between? Explain.

17 Young, a skateboarder of mass 70.0 kg, is riding along at 5.0 m s^{-1} east. Distracted by his mobile phone, he crashes head-first into a rigid metal pole and comes to a complete stop. Luckily, he is wearing his crash helmet that contains with compressible foam and the collision takes place over 0.350 s.

- **a** Calculate Young's momentum prior to the collision.
- **b** Calculate the average force of the pole applied to his head
- **c** If Young had forgotten his helmet, and the collision took place over 7.00 ms instead, what would be the average force applied to his head?
- **d** Apart from reducing the overall force applied to the head, how does a crash helmet protect its user in the event of a collision?

The following information applies to questions 18–20.

Jordy is playing softball and hits a ball with her softball bat. The force versus time graph for this interaction is shown below. The ball has a mass of 170 g.



18 Determine the magnitude of the change in momentum of the ball.

19 Determine the magnitude of the change in momentum of the bat.

20 Determine the magnitude of the change in velocity of the ball.

21 A pumpkin has a mass of 10 kg on Earth. What is its weight on Earth?

22 A skateboard has a weight of 20.6 N on Earth. What is its mass?

23 **a** is the mass of an 85 kg astronaut on the surface of Earth where g is 9.80 N kg^{-1}?

b What is the mass of an 85 kg astronaut on the surface of the Moon where g is 1.60 N kg^{-1}?

c What is the weight of an 85 kg astronaut on the surface of Mars where g is 3.60 N kg^{-1}?

24 Given the figures in question 19, order the weight of a 1 kg object from greatest weight to least weight when it is on the Moon, on Mars and on Earth.

CHAPTER 09

Work, energy and power

Throughout this chapter you will be learn about the common thread of energy conversion that is present in so many daily activities, as well as some more extreme activities. Your own personal energy stores are burnt up climbing steps or running to catch a bus. In more thrill-seeking adventures, such as bungee jumping, gravitational potential energy is converted into kinetic and elastic potential energy. Even jumping from a plane, the laws of physics cannot be switched off.

At the end of this chapter, you will be able to define and use the terms *work*, *energy* and *power*. You will use force–displacement graphs to determine the amount of work done.

Science Understanding

- Energy is conserved in isolated systems and is transferred from one object to another when a force is applied over a distance; this causes work to be done and changes the kinetic (E_k) and/or potential (E_p) energy of objects
  This includes applying the relationships

  $$E_k = \frac{1}{2}mv^2,\ E_p = mg\Delta h$$

  $$W = Fs,\ W = \Delta E$$

- Collisions may be elastic and inelastic; kinetic energy is conserved in elastic collisions.
  This includes applying the relationship

  $$\sum \frac{1}{2}mv^2{}_{\text{before}} = \sum \frac{1}{2}mv^2{}_{\text{after}}$$

- Power is the rate of doing work or transferring energy.
  This includes applying the relationship

  $$P = \frac{W}{t} = \frac{\Delta E}{t} = Fv_{av}$$

9.1 Energy and work

The words 'energy' and 'work' are commonly used to describe a variety of everyday situations. However, these words take on quite specific definitions when used in a scientific context. They are two of the most important concepts in physics, allowing physicists to explain phenomena on a range of scales from collisions of subatomic particles to the interactions of galaxies.

ENERGY

Energy is the capacity to cause change. A moving car has the capacity to cause a change if it collides with something else. Similarly, a heavy weight lifted by a crane has the capacity to cause a change if it is dropped. Energy is a scalar quantity; it has magnitude but not direction.

FIGURE 9.1.1 A moving car has kinetic energy.

There are many different forms of energy. **Mechanical energy** is defined as the energy that a body possesses due to its position or motion. This category of energy can be broadly classified into two groups: kinetic energy and potential energy.

Kinetic energy is energy associated with motion. Any moving object, like the moving car in Figure 9.1.1, has kinetic energy. In some forms of kinetic energy, the moving objects are not easily visible. An example of this is thermal energy, which is a type of kinetic energy related to the movement of particles. Table 9.1.1 lists some different types of kinetic energy and their associated moving objects.

TABLE 9.1.1 Types of kinetic energy and their associated moving objects.

Type of kinetic energy	Moving objects
translational kinetic	objects moving in a straight line
rotational kinetic	rotating objects
thermal	atoms, ions or molecules
sound	air molecules

Potential energy is energy associated with the position of objects relative to one another or within fields. For example, an object suspended by a crane has **gravitational potential energy** because of its position in the Earth's gravitational field. Some examples of potential energy are listed in Table 9.1.2.

TABLE 9.1.2 Types of potential energy and their causes.

Type of potential energy	Cause
gravitational	gravitational fields
chemical	relative positions of atoms
magnetic	magnetic fields
nuclear	forces within the nucleus of an atom
elastic	attractive forces between atoms

Unit of energy

The SI unit for energy, the joule (J), is named after the English scientist James Prescott Joule. He was the first person to show that kinetic energy could be converted into heat energy. The energy represented by 1 J is the equivalent to the energy needed to lift a 1 kg mass (e.g. 1 L of milk) through a height of 0.1 m or 10 cm. More commonly, scientists work in units of kilojoules (1 kJ = 1000 J) or even megajoules (1 MJ = 1000000 J).

WORK

Although in everyday life the word 'work' can take on a variety of meanings, in a scientific context work has a very specific meaning. In physics, when a force acts on an object and causes energy to be transferred or transformed, then work is being done on the object. For example, if a weightlifter applies a force to a barbell to lift it, then work has been done on the barbell; chemical energy within the weightlifter's body has been transformed into the gravitational potential energy of the barbell (Figure 9.1.2).

FIGURE 9.1.2 As a weightlifter lifts a barbell, chemical energy is transformed into gravitational potential energy.

Quantifying work

Work causes a change in energy, i.e. $W = \Delta E$.

More specifically, work is defined as the product of the force causing the energy change and the displacement of the object in the direction of the force during the energy change:

$W = Fs$
where W is work (J)
F is force (N)
s is the displacement in the direction of the force (m).

Since work corresponds to a change in energy, the SI unit of work is also the joule. The definition of work allows us to find a value for a joule in terms of other SI units.

Since $W = Fs$, $1\,\text{J} = 1\,\text{N} \times 1\,\text{m} = 1\,\text{N}\,\text{m}$.

A joule is equal to a newton-metre, that is, a force of 1 N acting over a distance of 1 m does 1 J of work.

Using the definition of a newton:

$1\,\text{J} = 1\,\text{N} \times 1\,\text{m} = 1\,\text{kg m s}^{-2} \times 1\,\text{m} = 1\,\text{kg m}^2\,\text{s}^{-2}$

This defines a joule in terms of fundamental units.

Although both force and displacement are vectors, work is a scalar unit. So, like energy, work has no direction.

PHYSICSFILE

Units of energy

A number of non-SI units for energy are still in use. When talking about the energy content of food, it is common to use a unit called a calorie (cal) (Figure 9.1.3). One calorie is defined as the amount of heat required to increase the temperature of 1 g of water by 1°C. This equates to 4.2 J.

Nutrition Facts

Serving Size 5 oz. (144g)
Servings Per Container 4

Amount Per Serving

Calories 310 **Calories** from Fat 100

	% Daily Value*
Total Fat 15g	**21%**
Saturated Fat 2.6g	**17%**
Trans Fat 1g	
Cholesterol 118mg	**39%**
Sodium 560mg	**28%**
Total Carbohydrate 12g	**4%**
Dietary Fiber 1g	**4%**
Sugars 1g	
Protein 24g	

Vitamin A 1% • **Vitamin C** 2%
Calcium 2% • **Iron** 5%

*Percent Daily Values are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs:

	Calories	2,000	2,500
Total Fat	Less Than	65g	80g
Saturated Fat	Less Than	20g	25g
Cholesterol	Less Than	300mg	300mg
Sodium	Less Than	2,400mg	2,400mg
Total Carbohydrate		300g	375g
Dietary Fiber		25g	30g

Calories per gram:
Fat 9 • Carbohydrate 4 • Protein 4

FIGURE 9.1.3 The amount of energy in a serving of food is often measured in calories.

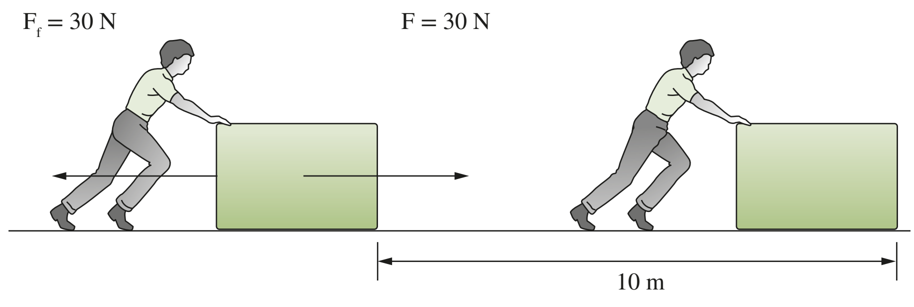
Electrical energy in the home is often measured in kilowatt-hours (kW h). A kilowatt-hour is a very large unit of energy:
1 kW h = 3 600 000 J or 3.6 MJ.

Another, not-so-common unit of energy is the erg (from the Greek word *ergon* for energy). An erg is a very small unit of energy: 1 erg = 10^{-7} J.

Worked example 9.1.1

CALCULATING WORK

A person pushes a heavy box along the ground for 10 m with a horizontal force of 30 N. Calculate the amount of work done.



Thinking	Working
Recall the definition of work.	$W = Fs$
Substitute in the values for this situation.	$W = 30 \times 10$
Solve the problem, giving an answer with appropriate units.	$W = 300$ J

Worked example: Try yourself 9.1.1

CALCULATING WORK

A person pushes a heavy wardrobe from one room to another by applying a force of 50 N for a distance of 5 m. Calculate the amount of work done.

Work and friction

The energy change produced by work is not always obvious. Consider Worked example 9.1.1, where 300 J of work was done on a box when it was pushed 10 m. A number of energy outcomes are possible for this scenario.

- In an ideal situation, where there was no friction, all of this work would be transformed into kinetic energy and the box would end up with a higher velocity than before it was pushed.
- In most real situations, where there is friction between the box and the ground, some of the work done would become heat and sound due to friction and the rest would become kinetic energy.
- In the limiting situation, where the force applied is exactly equal to the friction, the box would slide at a constant speed. This means that its kinetic energy would not change, so all of the work done would be converted into heat and sound due to friction.

Changing the displacement of a body is dependent on overcoming the force of friction.

A force with no work

The mathematical definition of work has some unusual implications. One is that if a force is applied to an object but the object does not move, then no work is done on the object.

This appears counterintuitive, that is, it goes against what you would probably expect. An example of this is shown in Figure 9.1.4. While picking up a heavy box requires work, holding the box at a constant height does no work on the box.

Assuming the box has a weight of 100 N and that it is lifted from the ground to a height of 1.2 m, the work done lifting it would be: $W = Fs = 100 \times 1.2 = 120$ J. In this case, energy is being transformed from chemical energy inside the person's body into the gravitational potential energy of the box.

However, when the box is held at a constant height, the definition of work gives: $W = Fs = 100 \times 0 = 0$ J. So, no work is being done *on* the box. Although there would be energy transformations going on inside the person's body to keep their muscles working, the energy of the box does not change, therefore no work has been done on the box.

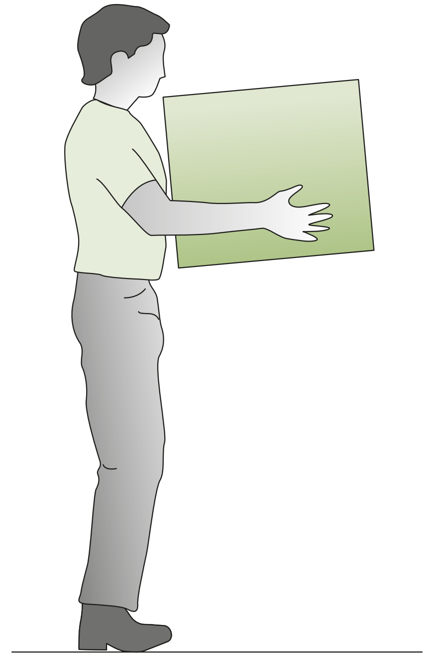


FIGURE 9.1.4 According to the definition of work, no work is done when a person holds a box at a constant height.

Work is done only if the net force causes a movement of one body in relation to other bodies.

Work and displacement at an angle

Sometimes, when a force is applied, the object does not move in the same direction as the force. For example, in Figure 9.1.5, when a person pushes a pram, the direction of the force is at an angle downwards, although the pram moves horizontally forwards.

FIGURE 9.1.5 When a person pushes a pram, the force is applied at an angle to the displacement of the pram.

In this case, only the *horizontal component* of the push contributes to the work being done on the pram. The vertical component of this force pushes the pram downwards and is balanced by the normal reaction force from the ground.

In situations like this, work can be calculated using the general equation:

$W = Fs \cos \theta$
where θ is the angle between the force, F, and the displacement, s.

EXTENSION

Resolving forces

In the situation of a person pushing a pram, the general equation $W = Fs$ applies. The person's push can be resolved into a vertical component, $F \sin \theta$, and a horizontal component, $F \cos \theta$, (Figure 9.1.6). Substituting the horizontal component into the general definition for work gives:

$$W = F \cos \theta \times s$$
$$= Fs \cos \theta$$

If the force applied is at right angles to the direction of displacement, then $\theta = 90°$, $\cos \theta = 0$ so $Fs \cos \theta = 0$, i.e. no work is done.

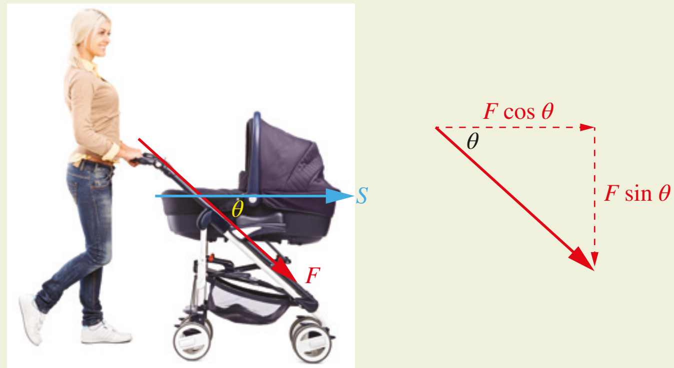
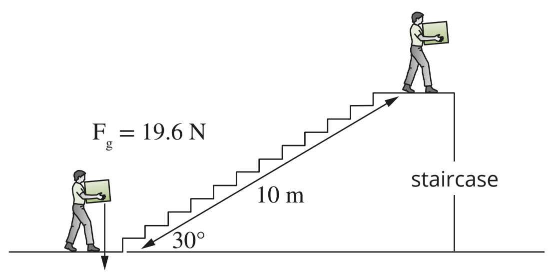


FIGURE 9.1.6 The force applied by the person pushing the pram can be resolved into a horizontal force and a vertical force.

Worked example 9.1.2

WORK WITH FORCE AND DISPLACEMENT AT AN ANGLE

A person carries a box weighing 19.6 N up a 10 m flight of stairs. Calculate the work done against gravity on the box.

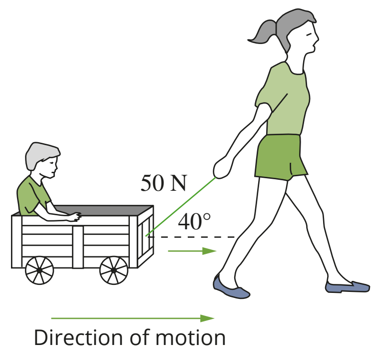


Thinking	Working
Determine values for F, s and θ. Note that the required component of the force is upwards, so the angle is not 30°. It is 90 – 30 = 60°.	Force applied to the box by the person: $F = 19.6$ N upwards Displacement: $s = 10$ m Angle between the force and displacement: $\theta = 60°$
Recall the work equation.	$W = Fs \cos \theta$
Substitute values into the work equation.	$W = 19.6 \times 10 \times \cos 60°$
State the answer with the correct units.	$W = 98$ J

Worked example: Try yourself 9.1.2

WORK WITH FORCE AND DISPLACEMENT AT AN ANGLE

A girl pulls her brother along in a trolley for a distance of 30 m, as shown. Calculate the work done on the box. Give your answer correct to three significant figures.



FORCE–DISPLACEMENT GRAPHS

As its name suggests, a force–displacement graph illustrates the way a force changes with displacement. For a situation where the force is constant, this graph is simple. For example, in Figure 9.1.7, the force–displacement graph for a person picking up a box is a flat horizontal line showing that the force applied to the box is constant throughout the lift.

In contrast, an **elastic** object such as a spring obeys a relationship known as Hooke's law. Hooke's law describes how, the more you stretch a spring, the greater the force required to keep stretching it. The force–displacement graph for a spring is also a straight line, but this line shows the direct relationship described by Hooke's law (Figure 9.1.8). (Note: sometimes, you will see force–displacement graphs for elastic objects labelled as force–extension graphs. In this context, the term extension is the same as displacement.)

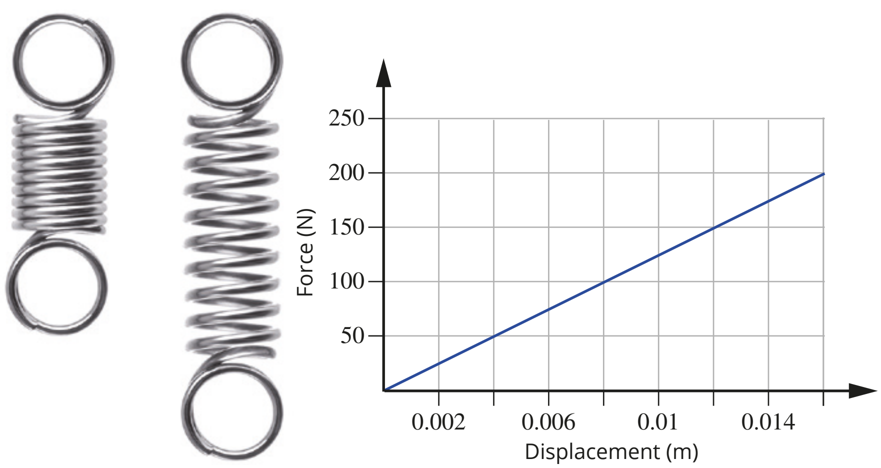


FIGURE 9.1.8 As a spring stretches, more force is required to keep stretching it. The force is proportional to the extension.

FIGURE 9.1.7 The force–displacement graph for a person picking up a box is a straight, horizontal line, indicating that the force applied to the box by the man is constant throughout the process.

Many everyday materials are only partially elastic. Their force–displacement graphs are relatively complex. For example, the force–displacement graph in Figure 9.1.9 for a sports shoe shows that the shoe is close to elastic for low displacements, but at high displacements the force is relatively constant.

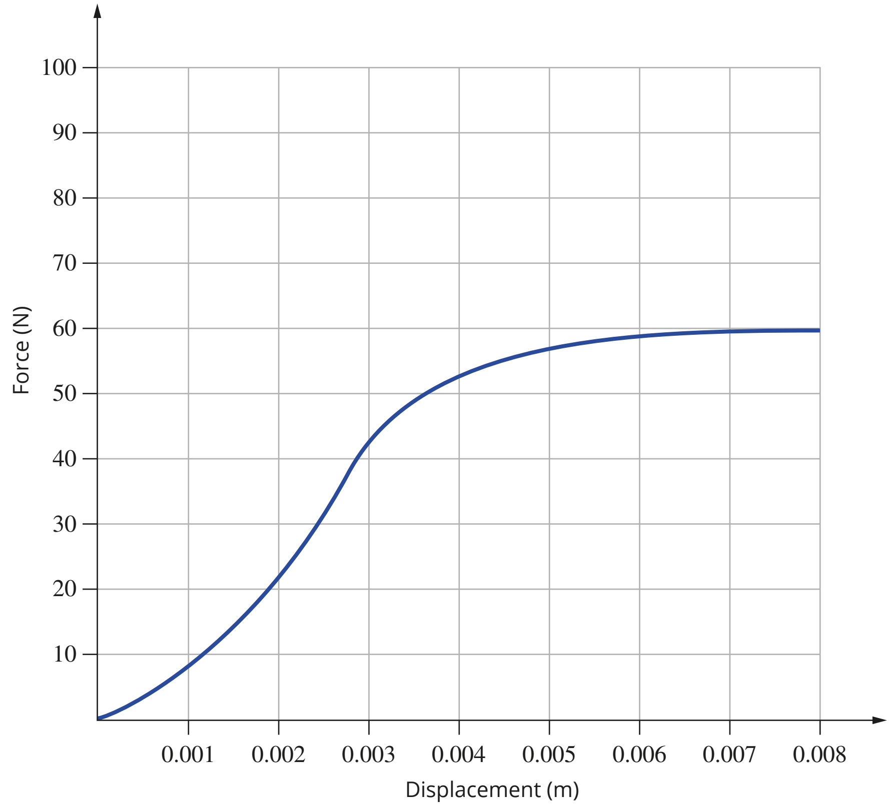


FIGURE 9.1.9 The force–displacement graph for a sports shoe is not a straight line; the change in force varies with how much the shoe has been stretched or compressed.

CALCULATING WORK FROM A FORCE–DISPLACEMENT GRAPH

When a force changes with displacement, the amount of work done by the force can be calculated from the area under its force–displacement graph.

For a constant force, this is very simple. Considering the earlier example of a person lifting a box, the area can be found by counting the number of 'force times distance' squares under the line. In the example in Figure 9.1.10 there are $6 \times 4 = 24$ of these squares. Since each square has an area of $5\,\text{N} \times 0.2\,\text{m} = 1\,\text{J}$, the total work done is 24 J. Alternatively, this area could be found by recognising that it is a rectangle and multiplying length by width to find the area. For Figure 9.1.10, this is $20\,\text{N} \times 1.2\,\text{m} = 24\,\text{J}$. Note that this second method is exactly the same as using the formula for work: $W = Fs = 20 \times 1.2 = 24\,\text{J}$. This relationship works in this case because the force is constant.

Similar strategies, either counting grid squares or calculating the area of the shape under the graph, can also be used when the force varies with the displacement.

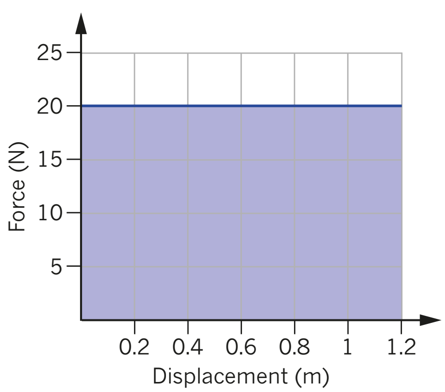
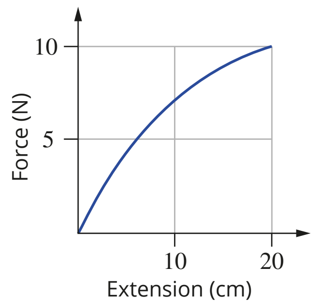


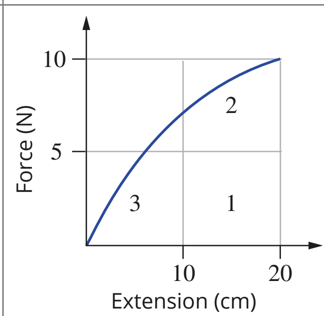
FIGURE 9.1.10 The area under a force–displacement graph gives the work done by the force.

Worked example 9.1.3

WORK FROM THE AREA UNDER A FORCE–DISPLACEMENT GRAPH

Use the force–extension graph for an elastic band to estimate how much work is done in stretching an elastic band 20 cm. Give your answer to the nearest 0.5 J.

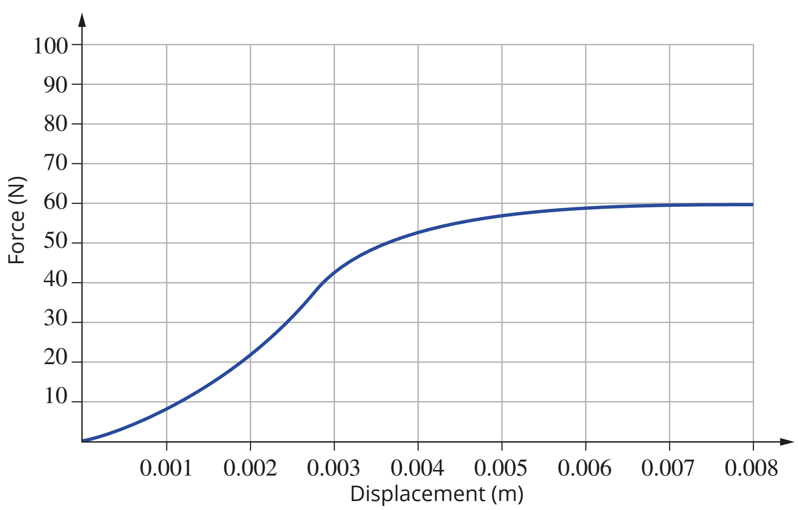


Thinking	Working
Calculate the work value of each grid square.	The dimensions of a grid square are: Force: 5 N, extension: 10 cm = 0.1 m Area of 1 square = 5 × 0.1 = 0.5 J
Count the number of grid squares under the curve. Only count grid squares that have more than half of their area under the curve. If the curve cuts a square exactly in half, count every second one.	 Number of squares = 3
Multiply the number of grid squares under the curve by the work value of each grid square.	$W = 3 \times 0.5\,\text{J} = 1.5\,\text{J}$

Worked example: Try yourself 9.1.3

WORK FROM THE AREA UNDER A FORCE–DISPLACEMENT GRAPH

While jogging, a person's shoes stretch by an average of 3 mm with each step. Use the force–displacement graph for a sports shoe to estimate how much work is done on the shoe with each step. Give your answer to the nearest 0.01 J.



9.1 Review

SUMMARY

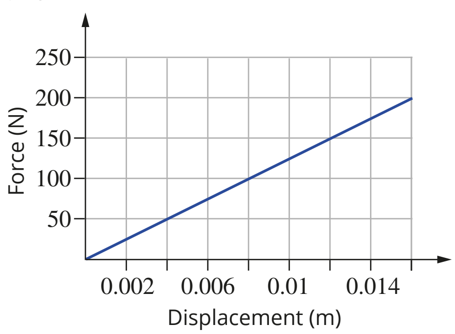
- Energy is the capacity to cause a change.
- Energy is conserved. It can be transferred or transformed, but not created or destroyed.
- There are many different forms of energy. These can be broadly classified as either kinetic (associated with movement) or potential (associated with the relative positions of objects).
- Work is done when energy is transferred or transformed.
- Work is done when a force causes an object to be displaced.
- Work is the product of force and displacement: $W = Fs$
- When a force produces no displacement, or when the force and displacement are at right angles to each other, no work is done.
- Work is equal to the area under a force–displacement graph.
- A straight horizontal line in a force–displacement graph represents a constant force.
- The relationship between force and displacement for a perfectly elastic object is represented as a straight diagonal line in a force–displacement graph.

KEY QUESTIONS

1 When accelerating at the beginning of a ride, a cyclist applies a force of 500 N for a distance of 20 m. What is the work done by the cyclist on the bike?

2 In the case of a person leaning on a solid brick wall, explain why no work is being done.

3 A spring with this force–displacement graph is stretched as shown. Using the formula for the area of a triangle, calculate the work done to stretch the spring by 0.015 m.

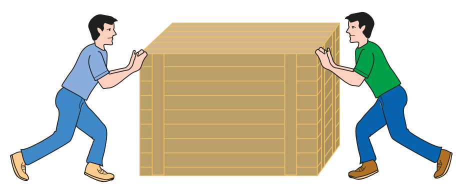


4 A cyclist does 2700 J of work when she rides her bike at a constant speed for 150 m. Calculate the average force the cyclist applies over this distance.

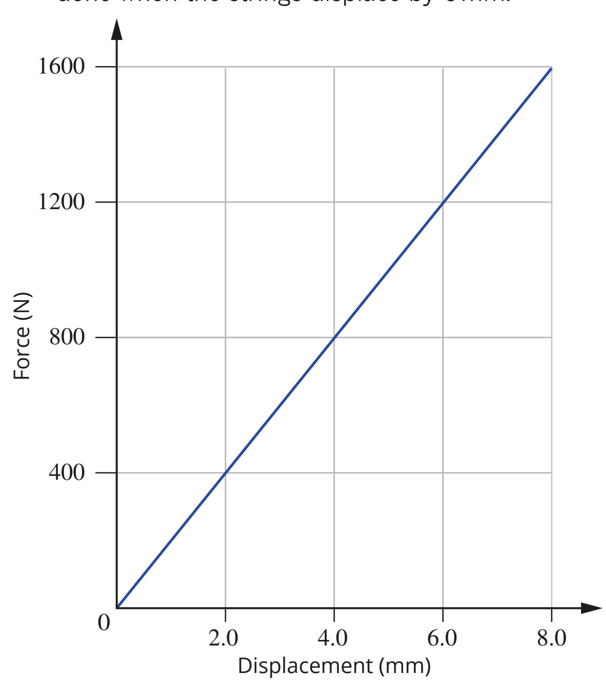
5 A rope at 40° to the horizontal is used to drag a heavy box along the ground for a distance of 5.0 m. Calculate the work done if the tension in the rope is 80 N. Give your answer correct to the nearest 10 J.

6 Explain why the equation $W = Fs$ cannot be used to calculate the work done in compressing a spring.

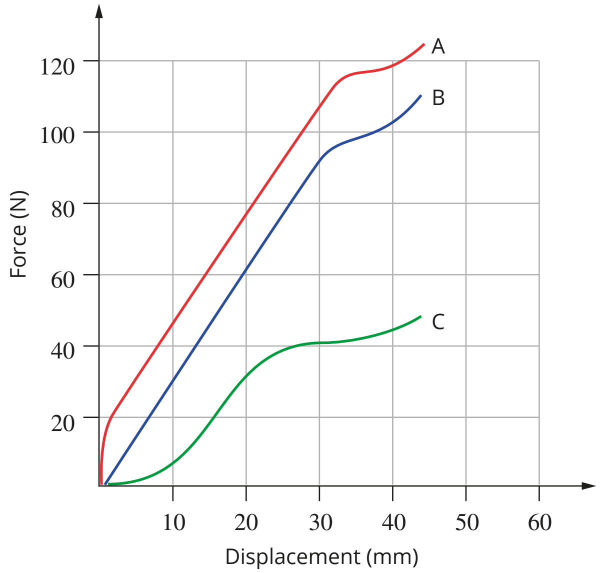
7 Two people push in opposite directions on a heavy box. One person applies 50 N of force, the other applies 40 N of force. There is 10 N of friction between the box and the floor which means that the box does not move. What is the work done by the person applying 50 N of force?



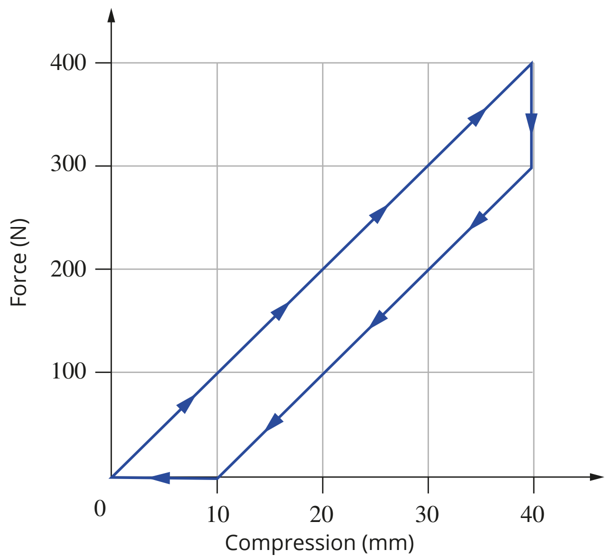
8 The strings of a graphite-head tennis racquet have the force–displacement graph shown. Calculate the work done when the strings displace by 6 mm.



9 Three different springs have the force–displacement graphs shown. Estimate the work done by stretching each of the springs by 40 mm. Give your answers correct to two significant figures.



10 The following diagram is a simplified representation of the forces acting on a basketball when it bounces. The upwards arrows show the force as the basketball compresses and the downwards arrows show the force as it rebounds.



a Calculate the work done on the basketball when it compresses by 40 mm.

b Calculate the work done by the ball as it decompresses from a compression of 40 mm.

c Explain why your answers to a and b are different.

9.2 Kinetic energy

Kinetic energy is energy that a body possesses due to its motion. Throwing a ball, rowing a canoe or launching a rocket ship—these all require energy as an inherent part of their motion.

Any object that moves, such as those shown in Figure 9.2.1, has kinetic energy. Many real-life energy interactions involve objects with kinetic energy. Some of these, like car collisions, have life-threatening implications. Hence, it is important to be able to quantify (find numerical values for) the kinetic energy of an object.

FIGURE 9.2.1 Any moving object, regardless of its size, has kinetic energy.

THE KINETIC ENERGY EQUATION

Kinetic energy is the energy of motion. It can be quantified by calculating the amount of work needed to give an object its velocity.

Consider the dynamics cart in Figure 9.2.2 of mass, m, starting at rest (i.e. $u = 0$). It is pushed with force, F, which acts while the cart undergoes a displacement, s, and gains a final velocity, v.

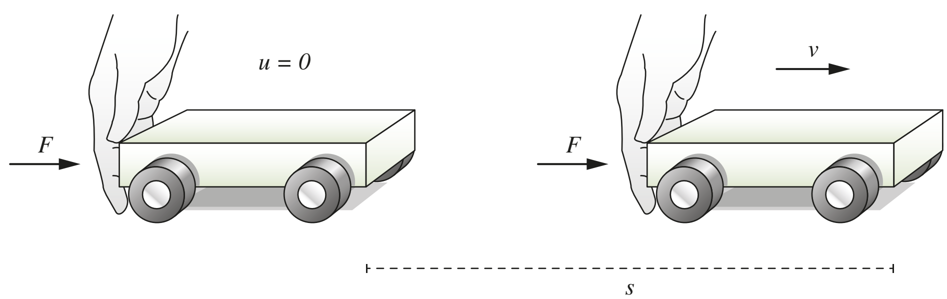


FIGURE 9.2.2 The kinetic energy of a dynamics cart can be calculated by considering the force, F, acting on it over a given displacement, s.

The work done by the force, W, causes a change in kinetic energy from its initial value $\frac{1}{2}mu^2$ to a new value of $\frac{1}{2}mv^2$.

The relationship between the work done and the change in kinetic energy can be written mathematically as:

$$W = \frac{1}{2}mv^2 - \frac{1}{2}mu^2$$

where W is work (J)
m is mass (kg)
u is initial velocity ($m\,s^{-1}$)
v is final velocity ($m\,s^{-1}$).
This equation is known as the 'work–energy theorem'.

In this situation, the cart was originally at rest (i.e. $u = 0$) so:

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

Assuming that no energy was lost as heat or noise and that all of the work is converted into kinetic energy, this equation gives us a mathematical definition for the kinetic energy of the cart in terms of its mass and velocity:

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

where E_k is kinetic energy (J)

EXTENSION

Expressing the amount of work

Considering the scenario described in Figure 9.2.2, the work done by the force is given by the equation $W = Fs$. The force causes the cart to accelerate according to Newton's second law, $F = ma$.

Rearranging the equation of motion $v^2 = u^2 - 2as$ gives:

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

Combining this with $F = ma$ means that the force acting on the cart can be given by the equation:

$$F = m\left(\frac{v^2 - u^2}{2s}\right)$$

This equation can be transposed to find an expression for the amount of work (Fs) done on the cart:

$$F = \frac{m}{2s}(v^2 - u^2)$$

$$Fs = \frac{m}{2}(v^2 - u^2)$$

$$Fs = \frac{1}{2}m(v^2 - u^2)$$

Since $W = Fs$:

$$W = \frac{1}{2}mv^2 - \frac{1}{2}mu^2$$

Worked example 9.2.1

CALCULATING KINETIC ENERGY

A car with a mass of 1200 kg is travelling at 90 km h^{-1}. Calculate its kinetic energy at this speed.	
Thinking	**Working**
Convert the car's speed to m s^{-1}.	$90\,\text{km h}^{-1} = \frac{90\,\text{km}}{1\,\text{h}} = \frac{90\,000\,\text{m}}{3600\,\text{s}}$ $= 25\,\text{m s}^{-1}$
Recall the equation for kinetic energy.	$E_k = \frac{1}{2}mv^2$
Substitute the values for this situation into the equation.	$E_k = \frac{1}{2} \times 1200 \times 25^2$
State the answer with appropriate units.	$E_k = 375\,000\,\text{J} = 375\,\text{kJ}$

Worked example: Try yourself 9.2.1

CALCULATING KINETIC ENERGY

A person crossing the street is walking at 5.0 km h^{-1}. If the person has a mass of 80 kg, calculate their kinetic energy. Give your answers correct to two significant figures.

APPLYING THE WORK-ENERGY THEOREM

The work–energy theorem can be seen as a definition for the *change* in kinetic energy produced by a force:

$$W = \frac{1}{2}mv^2 - \frac{1}{2}mu^2 = (E_k)_{final} - (E_k)_{initial} = \Delta E_k$$

Worked example 9.2.2

CALCULATING KINETIC ENERGY CHANGES

A 2 tonne truck travelling at 100 km h^{-1} slows to 80 km h^{-1} before turning a corner.

a Calculate the work done by the brakes to make this change. Give your answers correct to two significant figures.

Thinking	Working
Convert the values into SI units.	$u = 100\,\text{km h}^{-1} = \frac{100\,\text{km}}{1\,\text{h}} = \frac{100\,000\,\text{m}}{3600\,\text{s}}$ $= 28\,\text{m s}^{-1}$ $v = 80\,\text{km h}^{-1} = \frac{80\,\text{km}}{1\,\text{h}} = \frac{90\,000\,\text{m}}{3600\,\text{s}}$ $= 22\,\text{m s}^{-1}$ m = 2 tonne = 2000 kg
Recall the work–energy theorem.	$W = \frac{1}{2}mv^2 - \frac{1}{2}mu^2$
Substitute the values for this situation into the equation.	$w = \frac{1}{2}(2000 \times 22^2) - \frac{1}{2}(2000 \times 28^2)$
State the answer with appropriate units.	$W = -300\,000\,\text{J} = -300\,\text{kJ}$ Note: the negative value indicates that the work has caused the kinetic energy to decrease.

b If it takes 50 m for this deceleration to take place, calculate the average force applied by the truck's brakes.

Thinking	Working
Recall the definition of work.	$W = Fs$
Substitute the values for this situation into the equation. Note: The negative has been ignored as work is a scalar.	$300\,000\,\text{J} = F \times 50\,\text{m}$
Transpose the equation to find the answer.	$F = \frac{W}{s} = \frac{30\,0\,00}{50} = 6000\,\text{N}$

Worked example: Try yourself 9.2.2

CALCULATING KINETIC ENERGY CHANGES

As a bus with a mass of 10 tonnes approaches a school it slows from 60 km h^{-1} to 40 km h^{-1}.

a Calculate the work done by the brakes in the bus. Give your answers correct to two significant figures.

b The bus travels 40 m as it decelerates. Calculate the average force applied by the truck's brakes.

Notice that the definitions for kinetic energy and change in kinetic energy have been derived entirely from known concepts: the definition of work, Newton's second law and the equations of motion. This makes kinetic energy appear a redundant concept. However, using kinetic energy in calculations can often make analysis of physical interactions quicker and easier, particularly in situations where acceleration is not constant.

Worked example 9.2.3

CALCULATING SPEED FROM KINETIC ENERGY

The engine of a 1400 kg car can do 900 kJ of work in 10 s. Assuming all of this work is converted into kinetic energy, calculate the speed of the car after this time in km h^{-1}. Give your answer correct to two significant figures.

Thinking	Working
Recall the equation for kinetic energy.	$E_k = \frac{1}{2}mv^2$
Transpose the equation to make v the subject.	$v = \sqrt{\frac{2E_k}{m}}$
Substitute the values for this situation into the equation.	$v = \sqrt{\frac{2 \times 900 \times 10^3}{1400}} = 36\ \text{m s}^{-1}$
State the answer with appropriate units.	$v = 36 \times 3.6 = 130\ \text{km h}^{-1}$

Worked example: Try yourself 9.2.3

CALCULATING SPEED FROM KINETIC ENERGY

A 300 kg motorbike has 150 kJ of kinetic energy. Calculate the speed of the motorbike in km h^{-1}. Give your answer correct to two significant figures.

9.2 Review

SUMMARY

- All moving objects have kinetic energy.
- The kinetic energy of an object is equal to the work required to accelerate the object from rest to its final velocity.
- The kinetic energy of an object is given by the equation:
  - $E_k = \frac{1}{2}mv^2$
- The work–energy theorem defines work as *change* in kinetic energy:
  - $W = Fs = \frac{1}{2}mv^2 - \frac{1}{2}mu^2 = \Delta E_k$

KEY QUESTIONS

1 The mass of a motorbike together with its rider is 230 kg. If the motorbike is travelling at 80 km h^{-1}, calculate its kinetic energy.

2 A 1500 kg car is travelling at 17 m s^{-1}. How much work would its engine need to do to accelerate it to 28 m s^{-1}?

3 A cyclist has a mass of 72 kg and is riding a bicycle which has a mass of 9 kg. When riding at top speed on the bicycle, their kinetic energy is 5 kJ. Calculate the top speed of the cyclist in km h^{-1} to two significant figures.

4 By how much is kinetic energy increased when the mass of an object is doubled?

5 By how much is kinetic energy increased when the velocity of an object is tripled?

6 Lauren has eaten a 1.00 kJ biscuit. While she is out walking that day, she decides to speed up to a run to burn it all in one go. Assuming all of her work is converted into kinetic energy, and that she was previously going 0.500 m s^{-1} with a mass of 57.0 kg, how fast will she need to run in order to convert the biscuit into kinetic energy?

7 What is the difference in kinetic energy between a 1800 kg car travelling at 65 km h^{-1} and one travelling at 60 km h^{-1}? What might this suggest about the dangers of speeding?

9.3 Elastic and inelastic collisions

Collisions occur when two or more objects hit each other and exchange momentum and energy. These are important interactions. Energy and momentum can help us understand everyday collisions, for example between football players or between a bat and a ball, or more serious collisions, for example, car crashes (Figure 9.3.1). For physicists, studying collisions between subatomic particles may be the key to unlocking the secrets of the universe.

FIGURE 9.3.1 Collisions often occur in sport.

You have learned in Chapter 8 and the first part of this chapter that in all interactions, both energy and momentum are conserved. However, nature places no restrictions on the conversion of energy during collisions. During a collision or an interaction, kinetic energy can be entirely conserved, or it can be converted into other types of energy—such as heat energy, sound energy, and deformation of the objects.

ELASTIC AND INELASTIC COLLISIONS

Recall from the previous section that kinetic energy (E_k) is the energy of motion of a body. It is calculated using:

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

In perfectly **elastic collisions**, kinetic energy is transferred between objects, and no energy is transformed into heat or sound or deformation. In cases such as these, the relationship is stated as:

E_k (before) = E_k (after)

In Chapter 8 you saw how momentum is always conserved in a collision. The total energy is also always conserved in a closed system; however, in general, *kinetic* energy is not conserved, and energy is converted to heat or sound. Such collisions are called **inelastic collisions**.

Perfectly elastic collisions do not exist in everyday situations, but they do exist in the interactions between atoms and subatomic particles. A collision between two billiard balls or the spheres in a Newton's cradle is almost perfectly elastic because very little of their kinetic energy is transformed into heat and sound energy.

Collisions such as a bouncing basketball, a gymnast on a trampoline and a tennis ball being hit are moderately elastic, with about half the kinetic energy of the system being retained. Perfectly inelastic collisions are those in which the colliding bodies stick together after impact with no kinetic energy. Some car crashes, a collision between a meteorite and the Moon, and a collision involving two balls of plasticine could all be perfectly inelastic. In these collisions, most—and sometimes all—of the initial kinetic energy of the system is transformed into other forms of energy.

Worked example 9.3.1

ELASTIC OR INELASTIC COLLISION?

A car of mass 1.0×10^3 kg travelling west at $20\,\text{m}\,\text{s}^{-1}$ crashes into the rear of a stationary bus of mass 5.0×10^3 kg. The vehicles lock together on impact. Show calculations to test whether or not the collision is inelastic.

Thinking	Working
Use conservation of momentum to find the final velocity of the wreck.	$p_{\text{i car}} + p_{\text{i bus}} = p_{\text{f car+bus}}$ $mv_{\text{i car}} + mv_{\text{i bus}} = mv_{\text{f car+bus}}$ $(1.0 \times 10^3 \times 20) + (5.0 \times 10^3 \times 0)$ $= (1.0 \times 10^3 + 5.0 \times 10^3)v_{\text{f car+bus}}$ $v_{\text{f car+bus}} = \dfrac{2.0 \times 10^4}{6.0 \times 10^3}$ $v_{\text{f car+bus}} = 3.33\,\text{m}\,\text{s}^{-1}$
Calculate the initial kinetic energy before the collision for the bus and the car.	$E_{\text{ki bus}} = \frac{1}{2}mv^2$ $= \frac{1}{2}(5.0 \times 10^3) \times (0)^2$ $= 0\,\text{J}$ $E_{\text{ki car}} = \frac{1}{2}mv^2$ $= \frac{1}{2}(1.0 \times 10^3) \times (20)^2$ $= 2.0 \times 10^5\,\text{J}$ $E_{\text{ki}} = E_{\text{k i bus}} + E_{\text{k i car}}$ $= (0) + (2.0 \times 10^5)$ $= 2.0 \times 10^5\,\text{J}$
Calculate the final kinetic energy of the joined vehicles.	$E_{\text{k f}} = \frac{1}{2}mv^2$ $= \frac{1}{2}(5.0 \times 10^3 + 1.0 \times 10^3) \times (3.33)^2$ $= 3.3 \times 10^4\,\text{J}$
Compare the kinetic energy before and after the collision to determine whether or not the collision is elastic.	The kinetic energy after the collision is significantly less than the kinetic energy before. The missing energy has been transformed to heat, sound and deformation of the vehicles. Therefore, this collision is inelastic.

Worked example: Try yourself 9.3.1

ELASTIC OR INELASTIC COLLISION?

A 200 g snooker ball with initial velocity $9.0\,\text{m}\,\text{s}^{-1}$ to the right collides with a stationary snooker ball of mass 100 g. After the collision, both balls are moving to the right and the 200 g ball has a speed of $3.0\,\text{m}\,\text{s}^{-1}$. Show calculations to test whether or not the collision is inelastic.

EXTENSION

Top spin

Although perfectly elastic collisions only occur on the sub-microscopic scale (e.g. between atoms or subatomic particles), some everyday collisions are close enough to being elastic to make their outcomes predictable. For example, a collision between two balls in billiards or pool produces very little heat and sound and so is close to being elastic (Figure 9.3.2).

FIGURE 9.3.2 Pool, billiards and snooker are all games that involve nearly elastic collisions between balls.

This should mean that when a moving billiard ball strikes a stationary billiard ball, all of the momentum and kinetic energy should be transferred. The moving ball should stop and the stationary ball should move off at close to the initial velocity of the ball that struck it (Figure 9.3.3).

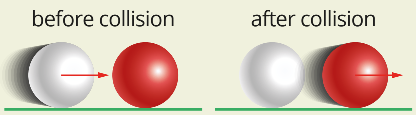


FIGURE 9.3.3 Since collisions between billiard balls are nearly elastic, both momentum and kinetic energy are transferred from the moving ball to the stationary ball.

However, in practice, a much greater variety of outcomes is possible from a billiard-ball collision. In fact, skilled billiards players are able to strike the cue ball (the white ball) in ways that can cause it to 'follow' (i.e. the moving ball continues rolling in its original direction after the collision; Figure 9.3.4) or to 'draw' (i.e. the moving ball rolls backwards after the collision; Figure 9.3.5).

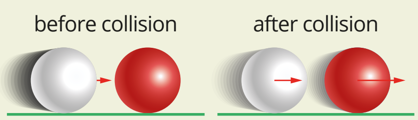


FIGURE 9.3.4 Billiards players can strike the cue ball so that, after the collision, it 'follows' the ball that it has struck.

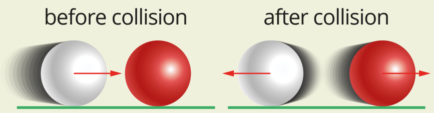


FIGURE 9.3.5 In a 'draw' shot, the player strikes the cue so that after the collision, it bounces back in the opposite direction.

These collisions appear to break the law of conservation of energy since kinetic energy seems to have either been created or destroyed. However, this is not the case. These sorts of shots are possible because there is more than one type of kinetic energy involved.

The energy that moves a ball from one spot on the table to another is known as translational kinetic energy. However, by hitting the cue ball slightly higher than its centre of mass, a billiards player can give the ball some top spin (Figure 9.3.6). This gives the ball some rotational kinetic energy, or energy of rotation, as well as translational kinetic energy. When the cue ball collides with another ball, some of the rotational kinetic energy is converted into translational kinetic energy, causing the cue ball to follow the ball it has struck.

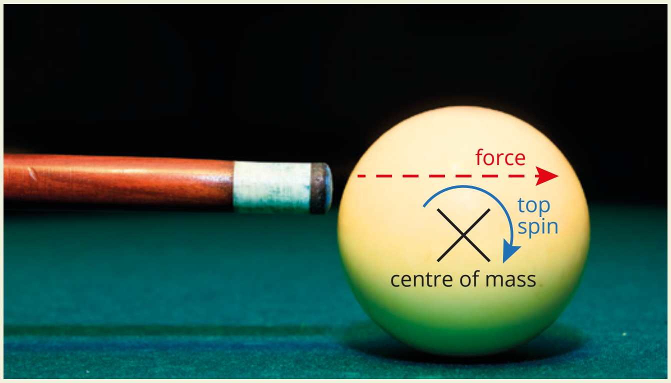


FIGURE 9.3.6 A billiards player can hit the cue ball slightly above the centre of mass of the ball to give it top spin.

Similarly, by striking the cue ball below its centre of mass, the ball can be given back spin. This rotational kinetic energy will cause the ball to move backwards after the collision.

Imparting spin to a ball is a common technique in many sports. The extra rotational energy that this gives the ball is not always obvious to other competitors and it can cause the ball to move or bounce in unexpected ways.

9.3 Review

SUMMARY

- Kinetic energy is the energy of motion of a body:
  $E_k = \frac{1}{2}mv^2$
- The total energy of an isolated system is always conserved.
- For perfectly elastic collisions, the kinetic energy before the collision is equal to the kinetic energy after the collision.

For inelastic collisions, the kinetic energy before the collision is greater than the kinetic energy after the collision; some of the initial kinetic energy is converted to heat or sound.

KEY QUESTIONS

1 Explain what the difference is between an elastic and an inelastic collision is and why an inelastic collision does not violate the law of conservation of energy.

2 A car of mass 1500 kg travelling east at $12.0\,\mathrm{m\,s^{-1}}$ crashes head on into a 2500 kg ute travelling at $16.0\,\mathrm{m\,s^{-1}}$ in the opposite direction. The vehicles lock together on impact.

a Calculate the velocity of the wreckage after the collision.

b Determine whether the collision is elastic or inelastic.

3 In a Newton's cradle, one ball is dropped so that it has a velocity of $1.50\,\mathrm{m\,s^{-1}}$ as it collides with the other four stationary balls. After the collision all five balls move together with the same speed. Use the law of conservation of momentum to determine the speed of the five balls after the collision and identify the collision as elastic or inelastic. Each ball has a mass of 40.0 g.

The following information applies to questions 4–7.

A 200 g toy truck with a springy bumper travelling at $0.300\,\mathrm{m\,s^{-1}}$ collides with a 100 g toy car travelling in the same direction at $0.200\,\mathrm{m\,s^{-1}}$. The car moves forward travelling at an increased speed of $0.300\,\mathrm{s^{-1}}$.

4 Calculate the speed of the truck after the collision.

5 Calculate the total kinetic energy of the system before the collision.

6 Calculate the total kinetic energy of the system after the collision.

7 Complete the following statements by selecting the appropriate option from those in brackets.

a The total kinetic energy before the collision is [more than/less than/equal to] the total kinetic energy after the collision.

b The kinetic energy of the system of toys [is/is not] conserved.

c The total energy of the system of toys [is/is not] conserved.

d The total momentum of the system of toys [is/is not] conserved.

e The collision [is/is not] perfectly elastic because [kinetic energy/total energy/momentum] is not conserved.

9.4 Gravitational potential energy

Gravitational potential is a measure of the amount of energy available to an object due to its position in a gravitational field. The gravitational potential energy of an object can be calculated from the amount of work that must be done against gravity to get the object into its position.

Consider the weightlifter lifting a barbell in Figure 9.4.1. Assuming that the bar is lifted at a constant speed, then the weightlifter must apply a lifting force equal to the force due to gravity on the barbell, F_g. The lifting force, F_l, is applied over a displacement, Δh, corresponding to the change in height of the barbell.

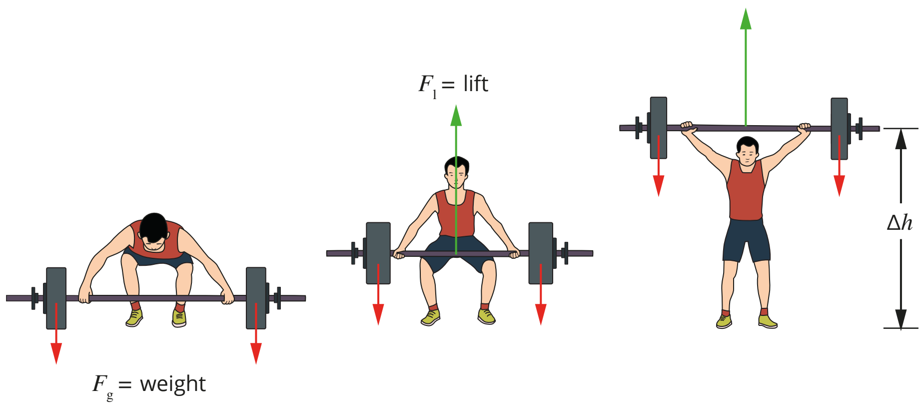


FIGURE 9.4.1 A weightlifter applies a constant force over a fixed distance to give the barbell gravitational potential energy.

The work done against gravity by the weightlifter is:

$W = Fs = F_g\Delta h$

Since the force due to gravity $F_g = mg$, the work done can be written as:

$W = mg\Delta h$

The work carried out in this example has resulted in the transformation of chemical energy within the weightlifter into gravitational potential energy. Since the change in gravitational potential energy of the barbell is:

$\Delta E_g = mg\Delta h$

Taking the ground as the point where the gravitational potential energy is zero (that is, $E_g = 0$) the gravitational potential energy of an object, due to the work done against a gravitational field, is given by:

$E_g = mg\Delta h$

where E_g is the gravitational potential energy (J)

m is the mass of the object (kg)

g is the gravitational field strength ($9.80\ \text{N kg}^{-1}$ on Earth)

Δh is the change in height of the object (m).

Note that the value of g is normally described as the acceleration due to gravity, measured in m s^{-2}. In this section, however, it is more appropriate to use g as the gravitational field strength. The value is the same, but the units are now N kg^{-1}. The two units are equivalent to each other. Gravitational field strength is used when considering factors affected by the gravitational field, while acceleration due to gravity is used for vertical motion in a gravitational field.

Worked example 9.4.1

CALCULATING GRAVITATIONAL POTENTIAL ENERGY

A weightlifter lifts a barbell which has a total mass of 80 kg from the floor to a height of 1.8 m above the ground. Calculate the gravitational potential energy of the barbell at this height. Give your answer correct to two significant figures.

Thinking	Working
Recall the formula for gravitational potential energy.	$E_g = mg\Delta h$
Substitute the values for this situation into the equation.	$E_g = 80 \times 9.80 \times 1.8$
State the answer with appropriate units and significant figures.	$E_g = 1411.2\ \text{J} = 1.4\ \text{kJ}$

Worked example: Try yourself 9.4.1

CALCULATING GRAVITATIONAL POTENTIAL ENERGY

A person doing their grocery shopping lifts a 5 kg grocery bag to a height of 30 cm. Calculate the gravitational potential energy of the grocery bag at this height. Give your answer correct to two significant figures.

GRAVITATIONAL POTENTIAL ENERGY AND REFERENCE LEVEL

When calculating gravitational potential energy, it is important to carefully define the level that corresponds to $E_g = 0$. Often this can be taken to be the ground or sea level, but the zero potential energy reference level is not always obvious.

It does not really matter which point is taken as the zero potential energy reference level, as long as the chosen point is used consistently throughout a particular problem (Figure 9.4.2). If objects move below the reference level, then their energies will become negative and should be interpreted accordingly.

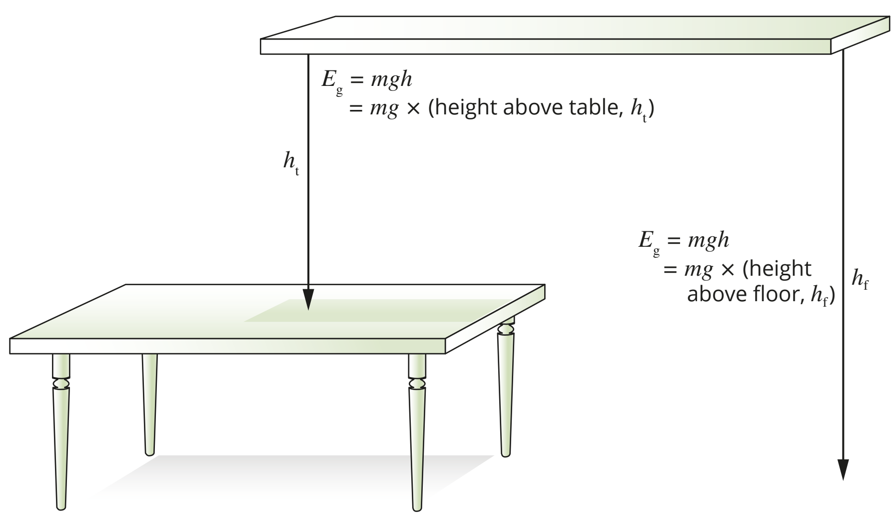


FIGURE 9.4.2 In this situation, the zero potential energy reference point could be taken as either the level of the table or the floor.

Worked example 9.4.2

CALCULATING GRAVITATIONAL POTENTIAL ENERGY RELATIVE TO A REFERENCE LEVEL

A weightlifter (m = 60 kg) lifts a 50 kg bar through a distance of 40 cm. Calculate the increase in gravitational potential energy of the bar with each lift. Use $g = 9.80\ \text{N kg}^{-1}$ and state your answer correct to three significant figures.	
Thinking	**Working**
Recall the formula for gravitational potential energy.	$E_g = mg\Delta h$
Identify the relevant values for this situation. Only the mass of the bar is being lifted (the weightlifter's mass is a distractor). Take the weightlifter's body as the zero potential energy level.	$m = 50\ \text{kg}$ $g = 9.80\ \text{N kg}^{-1}$ $\Delta h = 40\ \text{cm} = 0.4\ \text{m}$
Substitute the values for this situation into the equation.	$E_g = 50 \times 9.80 \times 0.4$
State the answer with appropriate units and significant figures.	$E_g = 196\ \text{J}$

Worked example: Try yourself 9.4.2

CALCULATING GRAVITATIONAL POTENTIAL ENERGY RELATIVE TO A REFERENCE LEVEL

A father picks up his baby from its bed. The baby has a mass of 6.0 kg and the mattress of the bed is 70 cm above the ground. When the father holds the baby in his arms, it is 125 cm off the ground. Calculate the increase in gravitational potential energy of the baby, taking g as $9.80\ \text{N kg}^{-1}$ and giving your answer correct to two significant figures.

EXTENSION

The high jump

Science has long been used in sport to help athletes gain a competitive edge. The concept of gravitational potential energy is of obvious importance to a high jumper. Clearly, the high jumper must do enough work in their jump to gain sufficient gravitational potential energy to clear the bar.

The modern high jump technique known as the Fosbury flop gets the high jumper to bend their body as they go over the bar. This is illustrated in Figure 9.4.3.

FIGURE 9.4.3 In the Fosbury flop technique, a high jumper must bend their body over the bar.

When the technique is correctly performed, most of the mass of the jumper (e.g. their head, arms and legs) is actually lower than the bar throughout the jump. In other words, the centre of mass of the jumper passes below the bar while their body bends over it, as shown in Figure 9.4.4.

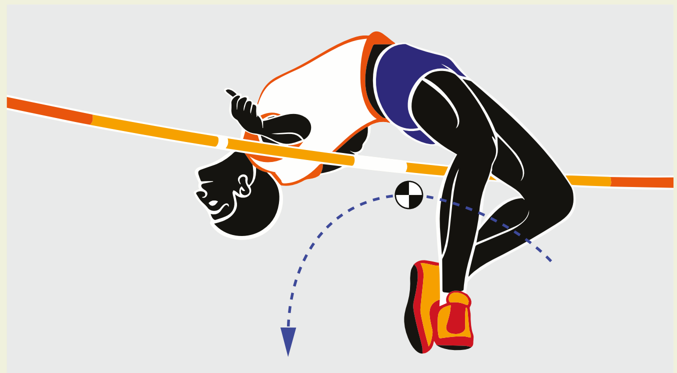


FIGURE 9.4.4 The path of the centre of mass of the high jumper (shown by the dashed curve) passes below the high-jump bar.

If the jumper's technique is right, they do not have to gain enough gravitational potential energy to lift all of their body higher than the bar for the time it takes to clear the bar. Without this technique, the world records for this event would probably be much lower than they currently are.

PHYSICSFILE

Newton's universal law of gravitation

The formula $E_g = mg\Delta h$ is based on the assumption that the Earth's gravitational field is constant. Newton's universal law of gravitation predicts that the Earth's gravitational field will decrease with altitude. However, this decrease only becomes significant many kilometres above the Earth's surface. For everyday purposes, the assumption of a constant gravitational field is valid.

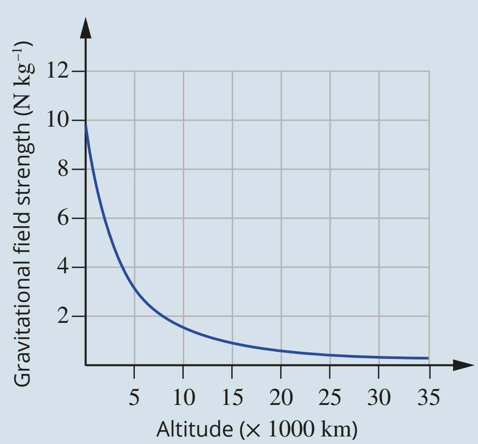


FIGURE 9.4.5 The Earth's gravitational field strength decreases with altitude.

9.4 Review

SUMMARY

- Gravitational potential energy is the energy an object has due to its position in a gravitational field.
- The gravitational potential energy of an object, E_g, is given by:
  $E_g = mg\Delta h$
- Gravitational potential energy is calculated relative to a zero potential energy reference level, usually the ground or sea level.

KEY QUESTIONS

1 A 57 g tennis ball is thrown 8.2 m into the air. Use $g = 9.80\ \text{N kg}^{-1}$.
  a Calculate the gravitational potential energy of the ball at the top of its flight.
  b Calculate the gravitational potential energy of the ball when it has fallen halfway back down to Earth.

2 When climbing Mount Everest ($h = 8848$ m), a mountain climber stops to rest at North Base Camp ($h = 5150$ m). If the mountain climber has a mass of 65.0 kg, how much gravitational potential energy will she gain in the final section of her climb (i.e. from North Base Camp to the summit)? For simplicity, assume that g is constant at $9.80\ \text{N kg}^{-1}$.

3 An astronaut visiting Mars has a mass of 90.0 kg. He climbs 60.0 m up a hill and his gravitational potential energy increases by 20.0 kJ. What is the strength of gravity on Mars?

4 While competing in high jump, Isabella (mass 50.0 kg, height 1.50 m) leaps with a jump of 550 J. She needs the centre of her body (0.75 m) to clear a 1.80 m bar. Neglecting air resistance, and assuming all of this energy is converted into gravitational potential energy, does she make it?

5 An eagle of mass 7.50 kg dives 150 m to catch a mouse. What is the eagle's change in potential energy?

6 If you were to lift a weight above your head and hold it still for several minutes, after the initial hoist is any additional work against gravity done? Why or why not?

9.5 Law of conservation of energy

In many situations, energy is transformed between kinetic and gravitational potential energy. For example, when a tennis ball bounces, as shown in Figure 9.5.1, much of its kinetic energy is converted into gravitational potential energy and then back into kinetic energy again.

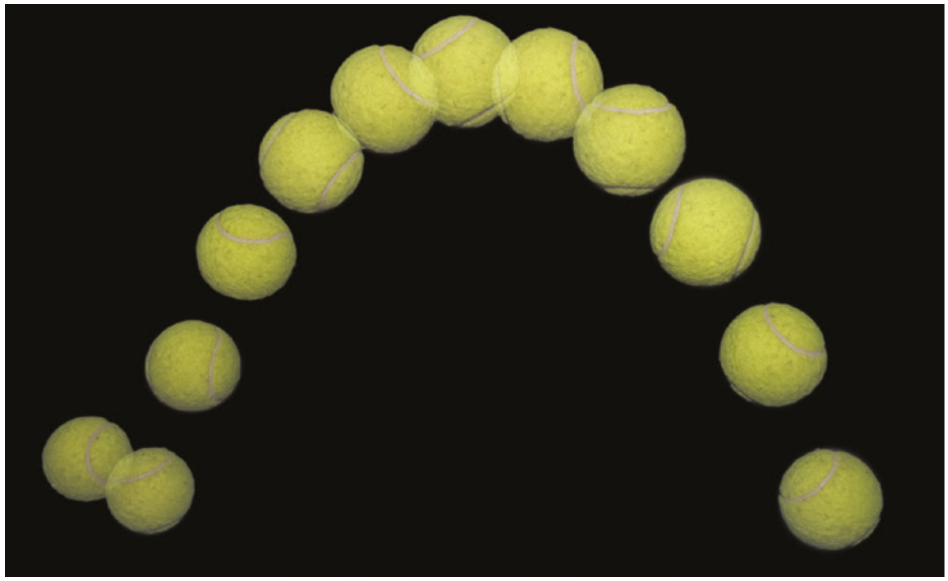


FIGURE 9.5.1 A bouncing tennis ball.

In analysing this type of situation, the concept of mechanical energy is useful. Mechanical energy is the sum of the potential energies available to an object and the kinetic energy of an object.

CONSERVATION OF ENERGY

Recall from Section 9.1 that energy comes in many forms, such as heat, light, sound, chemical energy and electrical energy. It is a scalar quantity and is measured in joules (J). Energy is also associated with the motion and position of an object, and this energy is called the mechanical energy of the object. In the motion problems of this chapter, moving objects have been described as having kinetic energy. An object can also have stored or gravitational potential energy because of its position. For instance, a building crane lifting a steel beam several stories has given the beam gravitational potential energy that could be disastrously converted to kinetic energy if the lifting chain were to break and the beam were to accelerate under gravity.

The changing of gravitational potential energy to kinetic energy is an illustration of the **law of conservation of energy**, a fundamental principle of nature. The law of conservation of energy states that energy is not created or destroyed, but can only change from one form to another, or in other words, **transform**. As the gravitational potential energy of the beam decreases, its kinetic energy increases. The total amount of mechanical energy remains constant.

While energy is not ever destroyed, it may be dissipated in forms that are not easily recoverable. For instance, the kinetic energy of a vehicle is reduced as it encounters friction, causing heating of the tyres, or in the deformation of the bodywork should it collide with another object. The mechanical energy before and after a collision is only the same under ideal conditions, but in many cases it is a useful approximation.

It is a common misconception to think that when energy is converted from a more obvious form (such as kinetic energy) that it has disappeared altogether. Consider the case of a car crashing into a tree. When the car stops, and the tree remains stationary, it may appear as though all energy has been lost, as nothing is moving.

In reality, all of the kinetic energy has been converted into more elusive forms—such as sound energy, heat energy, deformation energy, and so on. If you could collect all of this energy and compare it to the initial amount, it would be exactly equal to the initial energy. The energy would have been conserved.

MECHANICAL ENERGY

For a falling object, its mechanical energy is calculated from the sum of its kinetic and gravitational potential energies:

$$E_m = E_k + E_g = \frac{1}{2}mv^2 + mgh$$

This is a useful concept in situations where gravitational potential energy is converted into kinetic energy or vice versa. For example, consider a tennis ball with a mass of 60.0 g that is dropped from a height of 1.00 m (Figure 9.5.3). Initially, its total mechanical energy would comprise the kinetic energy, which would be 0 J, and the gravitational potential energy that is stored at this height (taking $g = 9.80\,\text{N}\,\text{kg}^{-1}$):

$$E_g = mgh = 0.0600 \times 9.80 \times 1.00 = 0.588\,\text{J}$$

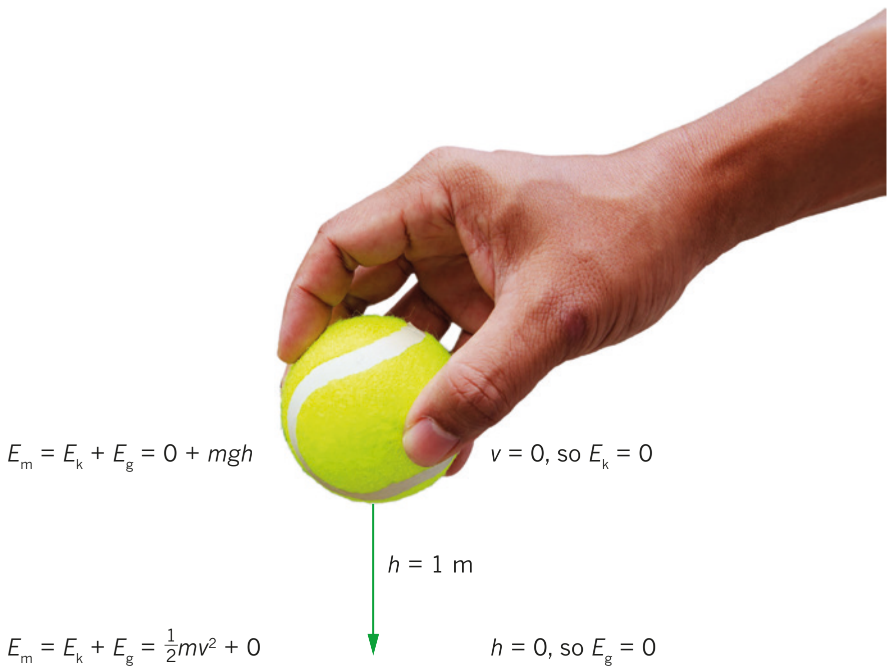


FIGURE 9.5.3 A falling tennis ball provides an example of conservation of mechanical energy.

At the instant the ball hits the ground, the total mechanical energy would comprise the gravitational potential energy available to it and the kinetic energy just prior to hitting the ground. The gravitational potential energy would be 0 J because the ball is at ground level. To calculate its kinetic energy, find the ball's velocity just before it hits the ground using one of the equations of motion:

$s = -1.00\,\text{m}$
$u = 0\,\text{m}\,\text{s}^{-1}$
$v = ?$
$a = -9.80\,\text{m}\,\text{s}^{-2}$
$t = ?$

$v^2 = u^2 + 2as$
$v^2 = 0^2 + 2(-9.80 \times -1.00)$
$v = \sqrt{19.6}$
$\quad = 4.43\,\text{m}\,\text{s}^{-1}$

Therefore the kinetic energy of the tennis ball just before it hits the ground is:

$$E_k = \frac{1}{2}mv^2 = \frac{1}{2} \times 0.0600 \times 4.43^2 = 0.588\,\text{J}$$

PHYSICSFILE

Noether's theorem

It might seem obvious in context now that energy is always conserved, but historically, it took centuries of debate and experiment to conclude just exactly *which* quantity is conserved. The conversion of kinetic energy into heat and sound made it difficult for scientists to keep track of the energy and justify that it wasn't, say, velocity which was conserved in general.

While the conservation of energy was a foundational law of thermodynamics, and many experiments had shown it as an empirical fact by the early twentieth century, it wasn't until 1915 that the brilliant German mathematician Emmy Noether provided a mathematical foundation and proof for general conserved quantities (such as energy and momentum) in theoretical physics. Her theorem has been described as one of the most important mathematical theorems ever proved in guiding the development of modern physics.

FIGURE 9.5.2 Emmy Noether, (1882–1935).

Notice that at both the top and the bottom of the 1.00 m fall, the mechanical energy is the same. At the top:

$E_m = E_k + E_g = 0 + 0.588 = 0.588\,J$

At the bottom:

$E_m = E_k + E_g = 0.588 + 0 = 0.588\,J$

In fact, mechanical energy is constant throughout the drop. Consider the tennis ball when it has fallen halfway to the ground. At this point, $h = 0.500\,m$ and $v = 3.13\,m\,s^{-1}$:

$$\begin{aligned} E_m &= E_k + E_g \\ &= \left(\frac{1}{2}\times 0.0600 \times 3.13^2\right) + (0.0600 \times 9.80 \times 0.500) \\ &= 0.294 + 0.294 \\ &= 0.588\,J \end{aligned}$$

Notice that, at this halfway point, the mechanical energy is evenly split between kinetic energy (0.294 J) and gravitational potential energy (0.294 J).

Throughout the drop, the mechanical energy has been conserved.

PHYSICSFILE

Mechanical energy of a ball falling through the air

In reality, as a ball drops through the air, a very small amount of its energy is transformed into heat and sound due to friction, and the ball won't quite reach a speed of $4.43\,m\,s^{-1}$ before it hits the ground. This means that mechanical energy is not entirely conserved. However, this small effect can be considered negligible for many falling objects.

The principle of conservation of mechanical energy states:
Given that, in a system of bodies, there are no other forms of energy except kinetic energy and potential energy, then the total mechanical energy of the system is constant.

Worked example 9.5.1

MECHANICAL ENERGY OF A FALLING OBJECT

A basketball with a mass of 600 g is dropped from a height of 1.2 m. Calculate its kinetic energy at the instant before it hits the ground.

Thinking	Working
Since the ball is dropped, its initial kinetic energy is zero.	$(E_k)_{initial} = 0\,J$
Calculate the initial gravitational potential energy of the ball.	$(E_g)_{initial} = mgh$ $= 0.600 \times 9.80 \times 1.2$ $= 7.1\,J$
Calculate the initial mechanical energy.	$(E_m)_{initial} = (E_k)_{initial} + (E_g)_{initial}$ $= 0 + 7.1$ $= 7.1\,J$
At the instant before the ball hits the ground, its gravitational potential energy is zero.	$(E_g)_{final} = 0\,J$
Mechanical energy is conserved in this situation.	$(E_m)_{initial} = (E_m)_{final} = (E_k)_{final} + (E_g)_{final}$ $\therefore\ 7.1 = (E_k)_{final} + 0$ $\therefore\ (E_k)_{final} = 7.1\,J$

Worked example: Try yourself 9.5.1

MECHANICAL ENERGY OF A FALLING OBJECT

A 6.8 kg bowling ball is dropped from a height of 0.75 m. Calculate its kinetic energy at the instant before it hits the ground.

USING MECHANICAL ENERGY TO CALCULATE VELOCITY

The speed of a falling object does not depend on its mass. This can be demonstrated using mechanical energy.

Consider an object with a mass, m, dropped from a height, h. At the moment it is dropped, its initial kinetic energy is zero. At the moment before it hits the ground, its final gravitational potential energy is zero. Therefore, using **conservation of mechanical energy**:

$$(E_m)_{\text{initial}} = (E_m)_{\text{final}}$$
$$(E_k)_{\text{initial}} + (E_g)_{\text{initial}} = (E_k)_{\text{final}} + (E_g)_{\text{final}}$$
$$0 + mgh = \frac{1}{2}mv^2 + 0$$
$$mgh = \frac{1}{2}mv^2$$
$$gh = \frac{1}{2}v^2$$
$$v^2 = 2gh$$
$$v = \sqrt{2gh}$$

This formula can be used to find the velocity of a falling object as it hits the ground. Note that the formula does not contain the mass of the falling object, so if air resistance is negligible, any object with any mass will have the same final velocity when it is dropped from the same height.

EXTENSION

Deriving a formula for velocity from the equations of linear motion

The velocity of a falling object formula can also be derived from the equations of linear motion as $a = g$. Consider an object dropped from a height, h, with negligible air resistance and an initial speed of $u = 0\,\text{m s}^{-1}$. Using the formula $v^2 = u^2 + 2as$:

$$\begin{aligned} v^2 &= u^2 + 2as \\ &= 0^2 + 2gh \\ &= 2gh \\ v &= \sqrt{2gh} \end{aligned}$$

This formula is equivalent to the result achieved using the conservation of mechanical energy.

Worked example 9.5.2

FINAL VELOCITY OF A FALLING OBJECT

A basketball with a mass of 600 g is dropped from a height of 1.2 m. Calculate its speed at the instant before it hits the ground.	
Thinking	**Working**
Recall the velocity of the falling object formula.	$v = \sqrt{2gh}$
Substitute the relevant values into the formula and solve.	$v = \sqrt{2 \times 9.80 \times 1.2}$ $= 4.8\ \text{m s}^{-1}$
Interpret the answer.	The basketball will be falling at $4.8\,\text{m s}^{-1}$ just before it hits the ground.

Worked example: Try yourself 9.5.2

FINAL VELOCITY OF A FALLING OBJECT

A 6.8 kg bowling ball is dropped from a height of 0.75 m. Calculate its speed at the instant before it hits the ground.

USING CONSERVATION OF MECHANICAL ENERGY IN COMPLEX SITUATIONS

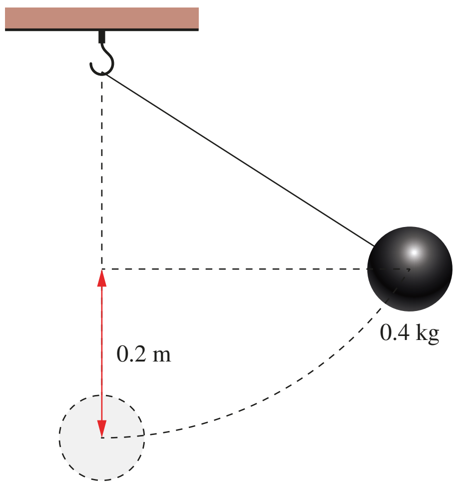


FIGURE 9.5.4 A falling pendulum provides an example of conservation of mechanical energy.

The concept of mechanical energy allows physicists to determine outcomes in non-linear situations where the equations of linear motion cannot be used. For example, consider a pendulum with a bob of mass 400 g displaced from its mean position such that its height has increased by 20 cm, as shown in Figure 9.5.4.

Since a falling pendulum involves gravitational potential energy being converted into kinetic energy, the conservation of mechanical energy applies to this situation. Therefore, the formula developed earlier for the velocity of a falling object can be used to find the velocity of the pendulum bob at its lowest point.

$$v = \sqrt{2gh} = \sqrt{2 \times 9.80 \times 0.2} = 2\ \text{m s}^{-1}$$

The speed of the pendulum bob will be $2\,\text{m s}^{-1}$ at its lowest point. However, unlike the falling tennis ball, the direction of the bob's motion will be horizontal instead of vertical at its lowest point. The equations of motion relate to linear motion and cannot be applied to this situation as the bob swings in a curved path.

Conservation of energy can also be used to analyse projectile motion – that is, when an object is thrown or fired into the air with some initial velocity. Since energy is not a vector, no vector analysis is required, even if the initial velocity is at an angle to the ground.

Worked example 9.5.3

USING MECHANICAL ENERGY TO ANALYSE PROJECTILE MOTION

A cricket ball (m = 140 g) is thrown upwards into the air at a speed of $15\,\text{m s}^{-1}$. Calculate the speed of the ball when it has reached a height of 8.0 m. Assume that the ball is thrown from a height of 1.5 m.

Thinking	Working
Recall the formula for mechanical energy.	$E_m = E_k + E_g = \frac{1}{2}mv^2 + mgh$
Substitute in the values for the ball as it is thrown.	$(E_m)_{initial} = \frac{1}{2}(0.14 \times 15^2) + (0.14 \times 9.80 \times 1.5)$ $= 18\,\text{J}$
Use conservation of mechanical energy to find an equation for the final speed.	$(E_m)_{final} = (E_k)_{final} + (E_g)_{final}$ $= \frac{1}{2}mv^2 + mgh$ $18 = \frac{1}{2}(0.14)v^2 + (0.14 \times 9.80 \times 8.0)$
Solve the equation algebraically to find the final speed.	$18 = 0.07v^2 + 11$ $7 = 0.07v^2$ $v^2 = \frac{7}{0.07}$ $= 100$ $v = \sqrt{100}$ $= 10\,\text{m s}^{-1}$
Interpret the answer.	The cricket ball will be moving at $10\,\text{m s}^{-1}$ when it reaches a height of 8.0 m.

Worked example: Try yourself 9.5.3

USING MECHANICAL ENERGY TO ANALYSE PROJECTILE MOTION

An arrow with a mass of 35 g is fired into the air at $80\,m\,s^{-1}$ from a height of 1.4 m. Calculate the speed of the arrow when it has reached a height of 30 m.

PHYSICS IN ACTION

Ballistics pendulum

The ballistics pendulum is an example of how the law of conservation of mechanical energy can be combined with an understanding of collisions to solve a practical problem. A ballistics pendulum is a device that can be used to measure the speed of a bullet fired from a gun or rifle. It consists of a block of wood hanging at a convenient height above the ground as shown in Figure 9.5.5.

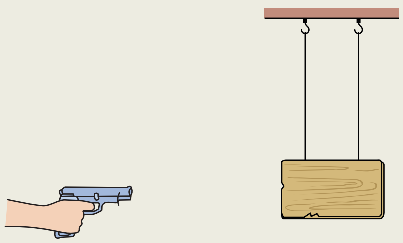


FIGURE 9.5.5 A ballistics pendulum combines an understanding of collisions and mechanical energy.

When a bullet is fired into the wooden block, an inelastic collision occurs. This means that much of the bullet's kinetic energy is converted into heat and sound and into changes made to the shape of the block. The conservation of mechanical energy does not apply for the impact of the bullet with the block.

However, the law of conservation of momentum still applies to the impact. This means that the block gains velocity from the bullet and it swings backwards and upwards as shown in Figure 9.5.6. By measuring the change in height of the block and the masses of the bullet and block, the initial speed of the bullet can be calculated. Note that conservation of mechanical energy does occur when the block swings backwards and upwards as no energy is converted into sound or heat during this part of its motion.

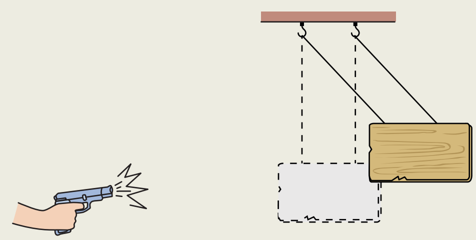


FIGURE 9.5.6 The change in height of a ballistics pendulum can be used to calculate the speed of the bullet fired into it.

LOSS OF MECHANICAL ENERGY

Mechanical energy is not conserved in every situation. For example, when a tennis ball bounces a number of times, each bounce is lower than the one before it, as shown in Figure 9.5.7.

While mechanical energy is largely conserved as the ball moves through the air, a significant amount of kinetic energy is transformed into heat and sound when the ball compresses and decompresses as it bounces. This means that the ball does not have as much kinetic energy when it leaves the ground as it did when it landed. Therefore, the gravitational potential energy it can achieve on the second bounce will be less than the gravitational potential energy it had initially, and so the second bounce is lower.

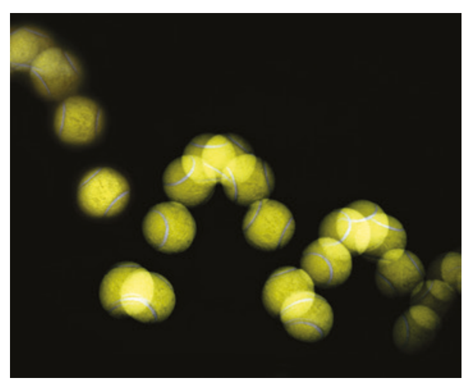


FIGURE 9.5.7 Mechanical energy is lost with each bounce of a tennis ball.

EXTENSION

Energy transformations in a bouncing ball

A bouncing ball involves forms of energy other than kinetic and gravitational potential energy. When the ball hits the ground, its gravitational potential energy is converted into elastic potential energy as it compresses. As the ball expands back to its original shape, some of the elastic potential energy is converted back into kinetic energy and some of it is converted into heat and sound. The amount of energy that is converted into heat and sound depends on the type of ball.

If you want the ball to reach a greater height than its original height, then instead of dropping the ball you could add to its energy by throwing it downwards with some velocity. Consider Figure 9.5.8a where a tennis ball ($m = 58\,\text{g}$) is thrown downwards at $4.0\,\text{m}\,\text{s}^{-1}$ from a height of 1 m.

Initially, the ball has 0.57 J of gravitational potential energy:

$$mgh = 0.058 \times 9.80 \times 1.0 = 0.57\,\text{J}$$

and 0.46 J of kinetic energy:

$$\frac{1}{2}mv^2 = \frac{1}{2}(0.058 \times 4^2) = 0.46\,\text{J}$$

By the time it reaches the ground, the gravitational potential energy has been transformed into kinetic energy, giving it a total of 1.04 J of kinetic energy (Figure 9.5.8b). This is converted into elastic potential energy of 1.03 J (Figure 9.5.8c).

If 0.28 J of energy are lost as heat and sound as the ball expands, then the ball will have just 0.75 J of kinetic energy when it leaves the ground (Figure 9.5.8d). This means that it will rebound to a height of 1.3 m (Figure 9.5.8e):

$$h = \frac{E_g}{mg} = \frac{0.75}{0.058 \times 9.80} = 1.3\text{ m}$$

Even though some energy has been 'lost' in the bounce, the initial kinetic energy of the ball means that it ends up slightly higher than where it started.

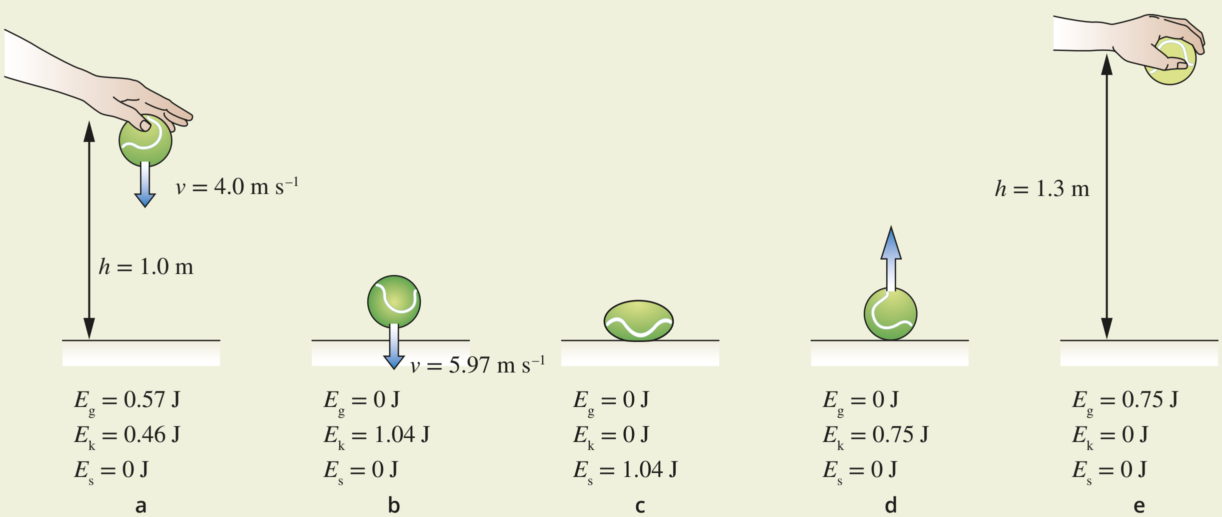


FIGURE 9.5.8 A tennis ball thrown downwards from a height.

EFFICIENCY OF ENERGY TRANSFORMATIONS

In the real world, energy transformations are never perfect—there is always some energy 'lost'. Because of this, for a system to continue operating (doing work), it must be constantly provided with energy. The percentage of energy that is effectively transformed by a device is called the **efficiency** of that device. A device operating at 45% efficiency is converting 45% of its supplied energy into the useful new form. The other 55% is 'lost' or transferred to the surroundings, usually as heat and/or sound. It is not truly lost since energy cannot be created or destroyed; rather, the form it becomes (heat and sound) is not useful.

The efficiency of a transfer from one energy form to another is expressed as:

$$\text{Efficiency } (\eta) = \frac{\text{(useful energy transferred)}}{\text{(total energy supplied)}} \times 100\%$$

$$= \frac{\text{energy output}}{\text{energy input}} \times 100\%$$

Table 9.5.1 shows the approximate efficiencies of some common objects.

TABLE 9.5.1 Efficiencies of some common objects or devices.

Device	Energy transfer	Efficiency (%)
electric motor	electric to kinetic	90
gas heater	chemical to thermal	75
incandescent light globe	electric to light	2
compact fluorescent light	electric to light	10
LED household light	electric to light	15
steam turbine	thermal to kinetic	45
coal-fired generator	chemical to electrical	30
high-efficiency solar cell	radiation to electrical	35
car engine	chemical to kinetic	25
open fireplace	chemical to thermal	15
human body	chemical to kinetic	25

Worked example 9.5.4

ENERGY EFFICIENCY

The energy input of a particular gas-fired power station is 1100 MJ. The electrical energy output is 300 MJ.

What is the efficiency of the power station?	
Thinking	**Working**
Recall the equation for efficiency. Substitute the given values into the equation.	output = 300 MJ input = 1100 MJ $\text{efficiency } (\eta) = \frac{\text{energy output}}{\text{energy input}} \times 100\%$ $\text{efficiency } (\eta) = \frac{300\text{ MJ}}{1100\text{ MJ}} \times 100\%$
Solve the equation.	efficiency = 27%

Worked example: Try yourself 9.5.4

ENERGY EFFICIENCY

An electric kettle uses 23.3 kJ of electrical energy as it boils a quantity of water. The efficiency of the kettle is 18%.

How much electrical energy is expended in actually boiling the water?

PHYSICSFILE

Coefficient of restitution

The ‘bounce of the ball’ is an important factor in many sports. Physicists describe the ‘bounciness’ of balls using a concept known as the coefficient of restitution (COR). COR depends on both the ball and the surface it is bouncing on. A tennis ball bouncing on grass has a different COR than one bouncing on clay. This is one reason why some tennis players prefer to play on some surfaces rather than others.

9.5 Review

SUMMARY

- The total energy of an isolated system is conserved in all circumstances.
- Mechanical energy is the sum of the potential and kinetic energies of an object.
- Mechanical energy is conserved in a falling object.
- Conservation of mechanical energy can be used to predict outcomes in a range of situations involving gravity and motion.
- The final velocity of an object falling with negligible air resistance from height *h* can be found using the equation $v = \sqrt{2gh}$
- When a ball bounces, some mechanical energy is transformed into heat and sound.
- The efficiency of an energy transfer from one form to another is:
  - Efficiency $(\eta) = \dfrac{\text{(useful energy transferred)}}{\text{(total energy supplied)}} \times 100\%$

$$= \frac{\text{energy output}}{\text{energy input}} \times 100\%$$

KEY QUESTIONS

1 A piano with a mass of 180 kg is pushed off the roof of a five-storey apartment block. The piano falls 3 m for each storey (i.e. a total of 15 m).

a Calculate the piano's kinetic energy as it hits the ground.

b Calculate the piano's kinetic energy as it passes the windows on the second floor, having fallen 10 m.

2 A tennis ball is dropped from the roof of a five-storey apartment block. The tennis ball falls 3 m for each storey (i.e. a total of 15 m).

a Calculate the tennis ball's speed as it hits the ground.

b Calculate the tennis ball's speed as it passes the windows on the second floor, having fallen 10 m.

3 A branch falls from a tree and hits the ground with a speed of 5.4 m s^{-1}. From what height did the branch fall?

4 A javelin with a mass of 800 g is thrown at an angle of inclination of 40°. It is released at a height of 1.45 m with a speed of 28.5 m s^{-1}.

a Calculate the javelin's initial mechanical energy.

b Calculate the speed of the javelin as it hits the ground.

5 A coal-fired generator has an efficiency of approximately 30%. If 2000 J of energy is supplied to the generator, how much is converted into electrical energy?

6 A rubber ball is dropped from a height of 1.5 m and loses 20% of its mechanical energy as it hits the ground. To what height will it rebound?

7 A girl rolls a ball up a frictionless hill. As the ball is travelling, does the kinetic energy of the ball increase or decrease? Does the gravitational potential energy increase or decrease? By what relative amount?

9.6 Power

In physics, it is instructive to look not only at absolute amounts of quantities, but also the rate at which these quantities change. Energy, in particular, is a quantity whose rate of change affects all aspects of society.

When considering energy changes, the rate at which work is done is often important. For example, if two cars have the same mass, then the amount of work required to accelerate each car from stationary to $100\,\text{km}\,\text{h}^{-1}$ will be the same. However, the fact that one car can do this more quickly than another may be an important consideration for some drivers when choosing which car to buy.

Physicists describe the rate at which work is done using the concept of **power**. Like work and energy, this is a word that takes on a specific meaning in a scientific context.

DEFINING POWER

Power is a measure of the rate at which work is done. Mathematically:

$$P = \frac{W}{\Delta t}$$

Recall that when work is done, energy is transferred or transformed. So the equation can also be written as:

$$P = \frac{\Delta E}{\Delta t}$$

where P is the power (W)

E is the energy transferred or transformed (J)

t is the time taken (s).

For example, a person running up a set of stairs does exactly the same amount of work as if they had walked up the stairs (i.e. $W = mgh$). However, the rate of energy change is faster for running up the stairs. Therefore, the runner is applying more power than the walker (Figure 9.6.1).

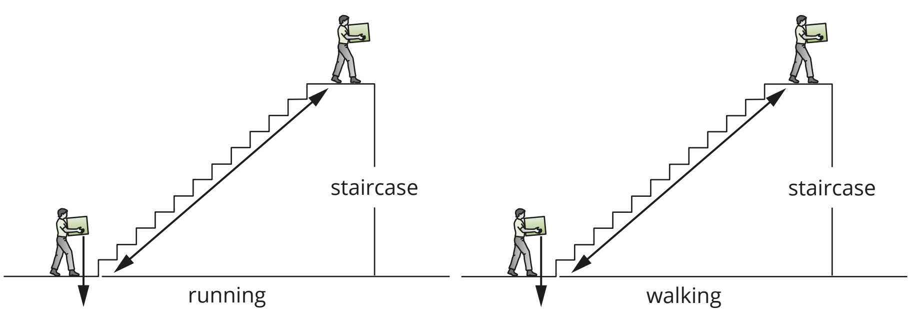


FIGURE 9.6.1 The runner and the walker both do the same amount of work, but the power output of the runner is higher than that of the walker.

Unit of power

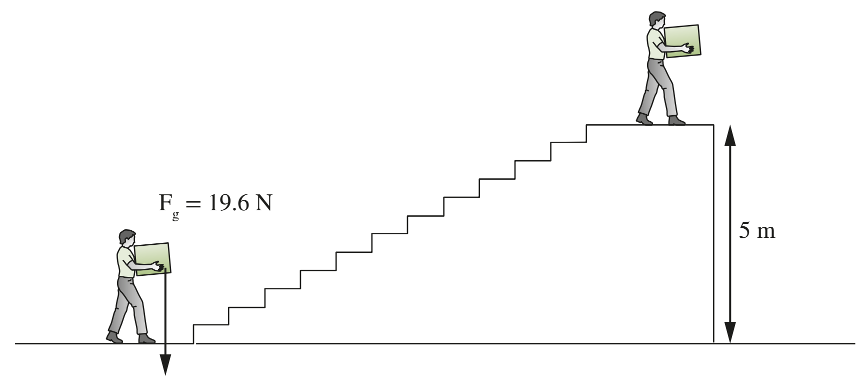
The unit of power is named after the Scottish engineer James Watt, who is most famous for inventing the steam engine. A watt (W) is defined as a rate of work of one joule per second; in other words:

$$1\text{W} = \frac{1\text{ J}}{1\text{ s}} = 1\text{ J s}^{-1}$$

Worked example 9.6.1

CALCULATING POWER

Calculate the power required to carry a box with a mass of 2 kg up a 5 m staircase in 20 s. (Use $g = 9.80\,\mathrm{m\,s^{-2}}$.)



Thinking	Working
Calculate the force applied.	$F_g = mg$ $= 2 \times 9.80$ $= 19.6\,\mathrm{N}$
Calculate the work done.	$W = Fs$ $= 19.6 \times 5$ $= 98\,\mathrm{J}$
Recall the formula for power.	$P = \frac{W}{\Delta t}$
Substitute the appropriate values into the formula.	$P = \frac{98}{20}$
Solve.	$P = 4.9\,\mathrm{W}$

Worked example: Try yourself 9.6.1

CALCULATING POWER

Calculate the power used by a weightlifter to lift a barbell which has a total mass of 50 kg from the floor to a height of 2.0 m above the ground in 1.4 s. (Use $g = 9.80\,\mathrm{m\,s^{-2}}$.)

PHYSICSFILE

Horsepower

James Watt was a Scottish inventor and engineer. He developed the concept of horsepower as a way to compare the output of steam engines with that of horses, which were the other major source of mechanical energy available at the time. Although the unit of one horsepower (1 hp) has had various definitions over time, the most commonly accepted value today is around 750 W. This is actually a significantly higher amount than an average horse can sustain over an extended period of time.

POWER, FORCE AND AVERAGE SPEED

In many everyday situations, a force is applied to an object to keep it moving at a constant speed—for example, pushing a wardrobe across a carpeted floor or driving a car at a constant speed. In these situations, the power being applied can be calculated directly from the force applied and the speed of the object.

Since $P = \frac{W}{\Delta t}$ and $W = Fs$, then:

$$P = \frac{Fs}{\Delta t} = F \times \frac{s}{\Delta t}$$

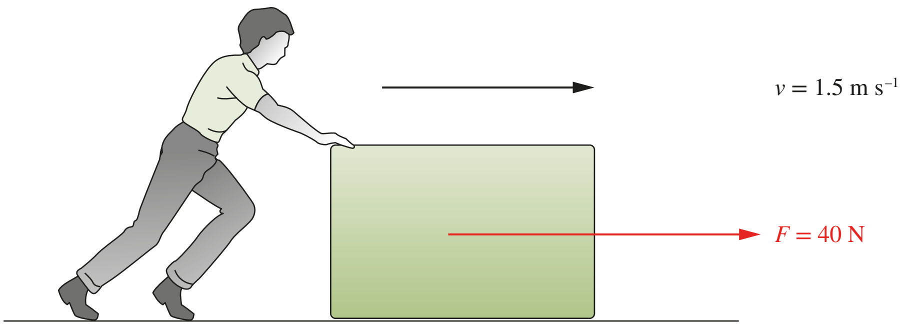
Since $\frac{s}{\Delta t}$ is the definition of average speed, the power equation can be written as:

$$P = Fv_{av}$$

Worked example 9.6.2

FORCE–VELOCITY FORMULATION OF POWER

A person pushes a heavy box along the ground at an average speed of $1.5\,\text{m s}^{-1}$ by applying a force of 40 N. What amount of power does the person exert on the box?



Thinking	Working
Recall the force–velocity formulation of the power equation.	$P = Fv_{av}$
Substitute the appropriate values into the formula.	$P = 40 \times 1.5$
Solve.	$P = 60\,\text{W}$

Worked example: Try yourself 9.6.2

FORCE–VELOCITY FORMULATION OF POWER

Calculate the power required to keep a car moving at an average speed of $22\,\text{m s}^{-1}$ if the force of friction (including air resistance) is 1200 N. Give your answer correct to three significant figures.

9.6 Review

SUMMARY

- Power is a measure of the rate at which work is done:
  $P = \frac{W}{\Delta t} = \frac{\Delta E}{\Delta t}$
- The power required to keep an object moving at a constant speed can be calculated from the product of the force applied and its average speed: $P = Fv_{av}$.

KEY QUESTIONS

1 A 1610 kg car accelerates from zero to $100\,\text{km h}^{-1}$ in 5.50 s. Calculate its average power output over this time.

2 A locomotive engine applies a force of 4 kN to keep a train moving at $20\,\text{m s}^{-1}$. Calculate the power output of the engine.

3 A 1700 kg car's engine uses 40 kW of power to maintain a constant speed of $80\,\text{km h}^{-1}$. Calculate the force being applied by its engine.

4 The motor of a crane has a maximum power output of 25 kW. At what average speed could it lift a concrete slab with a mass of 500 kg?

5 A box is pushed along a frictionless surface with a force of 15.0 N for 20.0 m in 10.0 s. Calculate the power output to push the box along.

6 A weightlifter lifts a 40.0 kg mass vertically upwards by 1.50 m. Given that this took 10.0 s to complete, what was the power generated by the weightlifter?

7 An 800.0 kg racecar's engine is capable of generating 120 kW of power. How long does it take the car to accelerate from 40.0 m s–1 to $55.0\,\text{m s}^{-1}$?

Chapter review

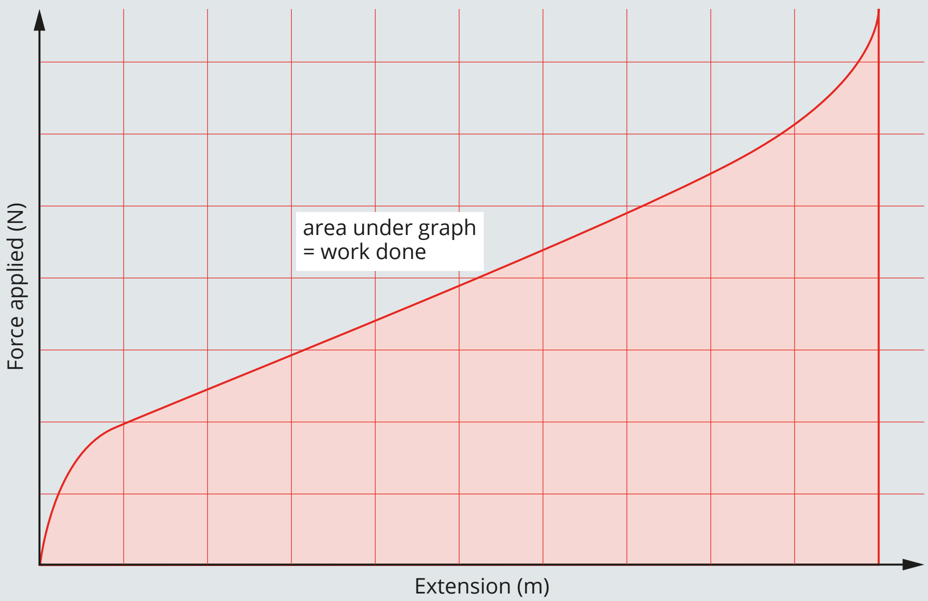
09

KEY TERMS

conservation of energy
conservation of mechanical energy
efficiency
elastic collision
gravitational potential energy
inelastic collision
kinetic energy
mechanical energy
power
transform
work

1 A car drives at a constant speed for 80 m. In order to overcome friction, its engine applies a force of 2000 N. Calculate the work done by the engine.

2 Estimate the total work shown in the following force–displacement graph, given that each grid square corresponds to 1 J.



3 A crane lifts a 200 kg load from the ground to a height of 30 m. What is the work done by the crane?

4 A person walks up a flight of 12 stairs. Each step is 240 mm long and 165 mm high. If the person has a mass of 60 kg and $g = 9.80\ \mathrm{m\,s^{-2}}$, what is the total amount of work done against gravity?

5 If 4000 J is used to lift a 50.0 kg object with a constant velocity, what is the theoretical maximum height to which the object can be raised?

6 A pram is pushed by the handle, which is at an angle of 35° to the horizontal (the direction of motion). If 1200 J of work is done pushing the pram 20 m, with what force was the pram pushed?

7 A cricket ball with a mass of 160 g is bowled with a speed of $150\ \mathrm{km\,h^{-1}}$. What is the kinetic energy of the cricket ball?

8 If a 1200 kg car has kinetic energy of 70 kJ, what is its speed?

9 The speed of an object is doubled. By how much does its kinetic energy increase?

10 A plumber (mass 88 kg) digs a ditch 40 cm deep. By how much does the plumber's gravitational potential energy change when he steps from the ground down into the ditch?

11 A football with a mass of 0.43 kg is kicked off the ground with a speed of $16\ \mathrm{m\,s^{-1}}$. How fast will it be going when it hits the crossbar, which is 2.44 m above the ground?

12 A white billiard ball moving with a speed of $5.00\ \mathrm{m\,s^{-1}}$ strikes a stationary red billiard ball. After the collision, the white ball continues in the same direction at a speed of $1.00\ \mathrm{m\,s^{-1}}$. The red ball rolls ahead of it (in the direction of the white ball's initial motion) with a speed of $4.00\ \mathrm{m\,s^{-1}}$. Each ball has a mass of 160g. Determine the energy of the objects both before and after the collision, and state whether it was elastic or inelastic.

13 A bullet of mass 10 g strikes a ballistics pendulum of mass 1.5 kg with speed *v* and becomes embedded in the pendulum. When the pendulum swings back, its height increases by 15 cm. For the questions that follow, assume that the initial gravitational potential energy of the pendulum was zero.

a What was the gravitational potential energy of the pendulum at the top of its swing?

b What was the kinetic energy of the pendulum when the bullet first embedded in it?

c What was the speed of the pendulum when it started to swing?

14 A crane can lift a load of 5 tonnes vertically through a distance of 20 m in 5 s. What is the power of the crane approximately?

15 A red Mini Cooper with a mass of 650 kg can accelerate from 0 to 100 km h^{-1} in 7.2 s. What is the average power output of the car over this time?

16 If the engine of a 1400 kg car uses 25 kW to maintain an average speed of 17 m s^{-1}, how much friction is acting on the car?

17 At the start of a 100 m race, a runner with a mass of 60 kg accelerates from a standing start to 8.0 m s^{-1} in a distance of 20 m.

a Calculate the work done by the runner's legs.

b Calculate the average force that the runner's legs apply over this distance.

18 When moving around on the Moon, astronauts find it easier to use a series of small jumps rather than to walk. If an astronaut (with a mass of 120 kg including equipment) jumps to a height of 10 cm on the Moon, where the gravitational field strength is 1.6 m s^{-2}, by roughly how much does his potential energy increase?

19 The efficiency of an appliance is known to be 80%. What energy was supplied if the output was 1250 J?

CHAPTER 10

The nature of waves

Have you ever watched ocean waves heading toward the shore? For many people, their first thought when encountering a topic called ‘waves’ is to picture a water wave moving across the surface of an ocean. The wave may be created by some kind of disturbance like the action of wind on water or a boat as it moves through the water.

Waves are everywhere. Sound, visible light, radio waves, waves in the string of an instrument, the wave of a hand, the ‘Mexican wave’ at a stadium and the recently discovered gravitational waves—all are waves or wave-like phenomena. Understanding the physics of waves provides a broad base upon which to build your understanding of the physical world.

Science as a Human Endeavour

Application of the wave model has enabled the visualisation of imaging techniques. These can include:

- medical applications, such as ultrasound
- geophysical exploration, such as seismology

Noise pollution comes from a variety of sources and is often amplified by walls, buildings and other built structures. Acoustic engineering, based on an understanding of the behaviour of sound waves, is used to reduce noise pollution. It focuses on absorbing sound waves or planning structures so that reflection and amplification do not occur.

Science Understanding

- waves are periodic oscillations that transfer energy from one point to another
- mechanical waves transfer energy through a medium; longitudinal and transverse waves are distinguished by the relationship between the directions of oscillation of particles relative to the direction of the wave velocity
- waves may be represented by displacement/time and displacement/distance wave diagrams and described in terms of relationships between measurable quantities, including period, amplitude, wavelength, frequency and velocity
  This includes applying the relationships
  $$v = f\lambda\ ,\ T = \frac{1}{f}$$
- the mechanical wave model can be used to explain phenomena related to reflection and refraction, including echoes and seismic phenomena
- the superposition of waves in a medium may lead to the formation of standing waves and interference phenomena, including standing waves in pipes and on stretched strings

This includes applying the relationships for:
strings attached at both ends and pipes open at both ends

$$\lambda_n = \frac{2l}{n} \text{ where } n \text{ is the harmonic, } n = 1, 2, 3, \ldots$$

and for pipes closed at one end

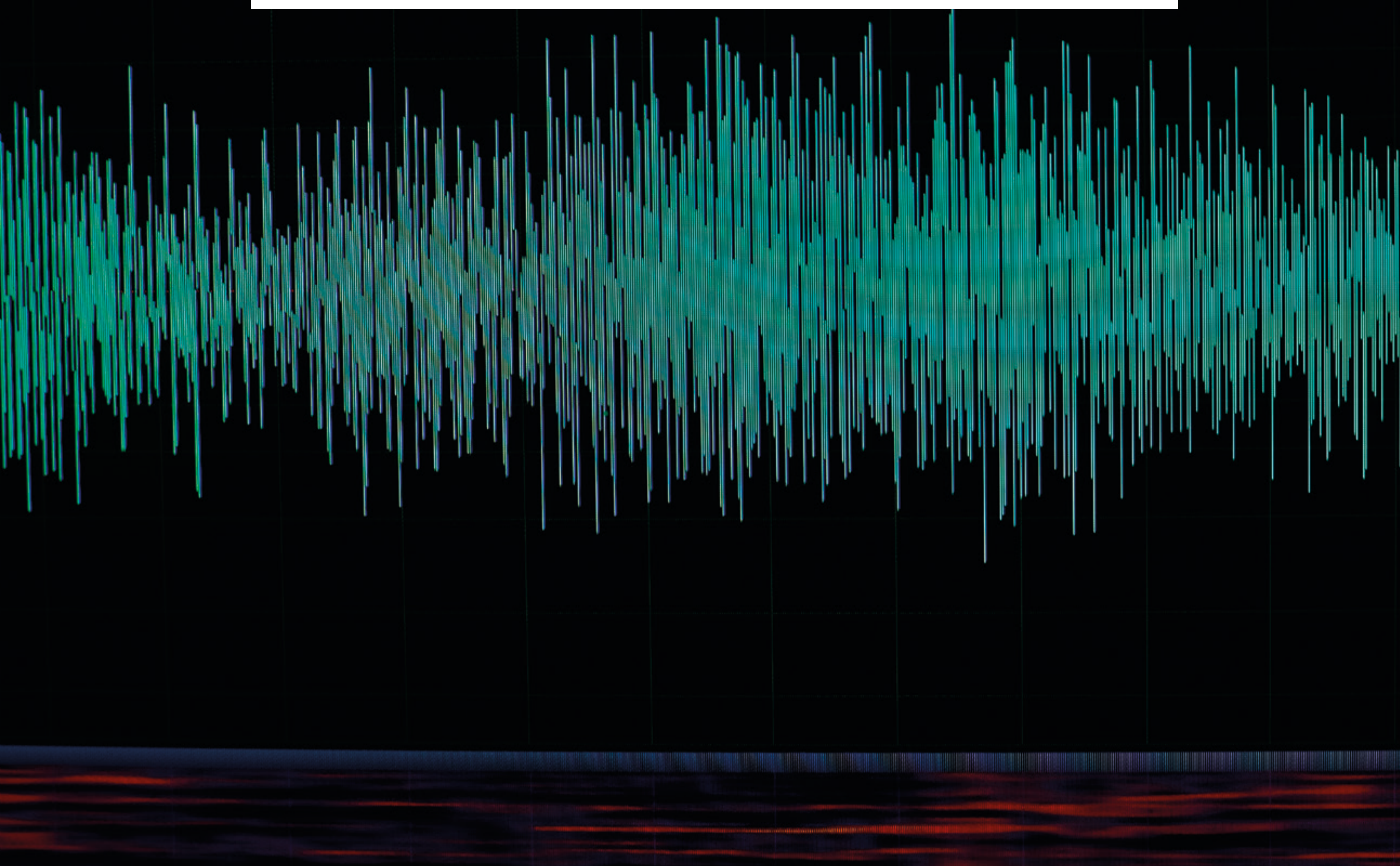
$$\lambda_\mathrm{n} = \frac{4l}{n} \text{ where n is the harmonic number, n} = 1, 3, 5\ldots$$

or

$$\lambda_{2n-1} = \frac{4l}{2n-1} \text{ where n} = 1, 2,3, \ldots \text{ and } 2n-1 \text{ is the harmonic number}$$

- a mechanical system resonates when it is driven at one of its natural frequencies of oscillation; energy is transferred efficiently into systems under these conditions
- the intensity of a wave decreases in an inverse square relationship with distance from a point source
  This includes applying the relationship

$$I \propto \frac{1}{r^2}$$



10.1 Longitudinal and transverse waves

Throw a stone into a pool or lake, and you will see circular waves form and move outwards from the source as ripples, as shown in Figure 10.1.1. Stretch a cord out on a table and wriggle one end back and forth across the table surface and another type of wave can be observed. Sound waves, water waves and waves in strings are all examples of **mechanical waves**. Mechanical waves need a medium to transmit energy. Electromagnetic waves, on the other hand, transmit energy through a vacuum. Mechanical waves are the focus of this chapter. Electromagnetic waves are discussed more fully in Year 12.

FIGURE 10.1.1 The ripples in a pond indicate a transfer of energy.

MECHANICAL WAVES

Watch a piece of driftwood, a leaf, or even a surfer, resting in the water as a smooth wave goes past. The object moves up and down but doesn't move forward with the wave. The movement of the object on the water reveals how the particles in the water move as the wave passes; that is, the particles in the water move up and down from an average position.

Any wave that needs a **medium** (such as water) through which to travel is called a mechanical wave. Mechanical waves can move over very large distances but the particles of the medium only have very limited movement.

In mechanical waves the particles of the matter vibrate back and forth or up and down about an average position, which transfers the energy from one place to another. For example, energy is given to an ocean wave by the action of the wind far out at sea. The energy is transported by waves to the shore but (except in the case of a tsunami event) most of the ocean water itself does not travel onto the shore.

A wave involves the transfer of energy without the net transfer of matter.

PULSES VERSUS PERIODIC WAVES

A single wave **pulse** can be formed by giving a slinky spring or rope a single up and down motion as shown in Figure 10.1.2a. As the hand pulls upwards, the adjacent parts of the slinky will also feel an upward force and begin to move upward. The source of the wave energy is the movement of the hand.

If the up and down motion is repeated, each successive section of the slinky will move up and down, moving the wave forward along the slinky as shown in Figure 10.1.2b. Connections between each loop of the slinky cause the wave to travel away from the source, carrying with it the energy from the source.

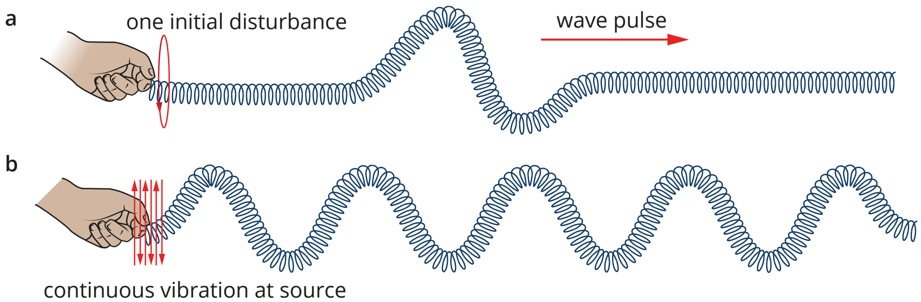


FIGURE 10.1.2 (a) A single wave pulse can be sent along a slinky by a single up and down motion. (b) A continuous or periodic wave is created by a regular, repeated movement of the hand.

In a continuous wave or periodic wave, continuous vibration of the source, such as that shown in Figure 10.1.2b, will cause the particles within the medium to **oscillate** about their average position in a regular, repetitive or periodic pattern. The source of any mechanical wave is this repeated motion or **vibration**. The energy from the vibration moves through the medium and constitutes a mechanical wave. Any wave that travels unimpeded through a medium is known as a **travelling wave**.

Transverse waves

When waves travel on water, or through a rope, spring or string, the particles within the medium can vibrate up and down in a direction perpendicular, or **transverse**, to the direction of motion of the wave energy (Figure 10.1.3). Such a wave is called a transverse wave. When the particles are displaced upwards from the average position, or resting position, they reach a maximum positive displacement called a **crest**. Particles below the average position fall to a maximum negative position called a **trough**.

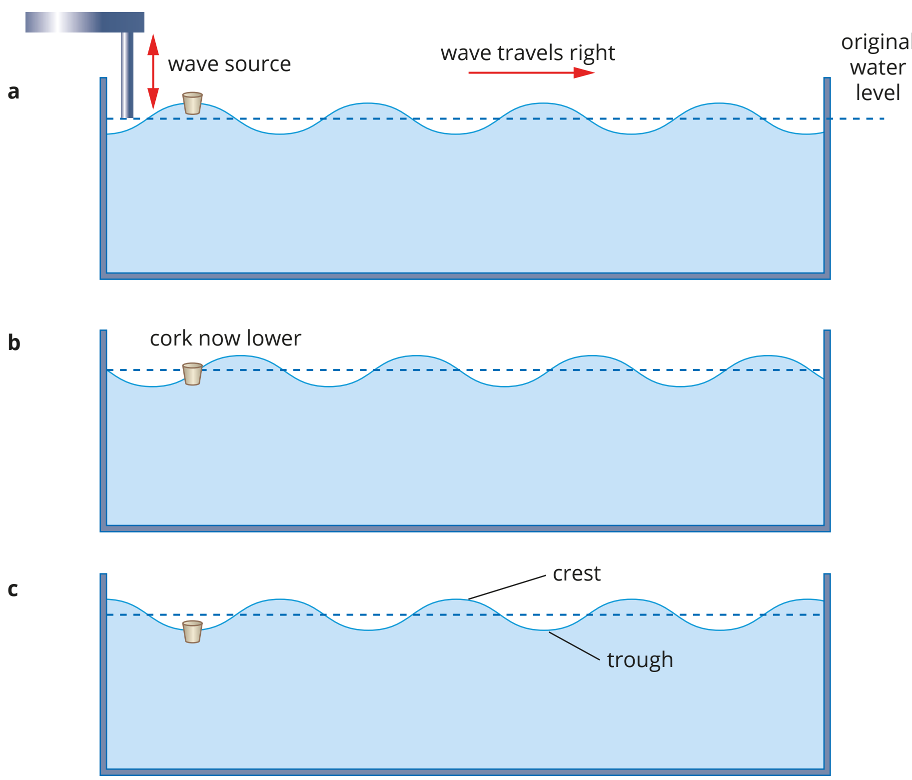


FIGURE 10.1.3 A continuous water wave moves to the right showing an example of a transverse wave. As it does so the up and down displacement of the particles transverse to the wave motion can be monitored using a cork. The cork simply moves up and down as the wave passes through it.

A transverse wave oscillates or vibrates in a direction perpendicular to the direction of the wave.

Longitudinal waves

In a **longitudinal** mechanical wave, the vibration of the particles within the medium are in the same direction, or parallel to, the direction of energy flow of the wave. You can demonstrate this type of wave with a slinky by moving your hand backwards and forwards in a line parallel to the length of the slinky, as shown in Figure 10.1.4a.

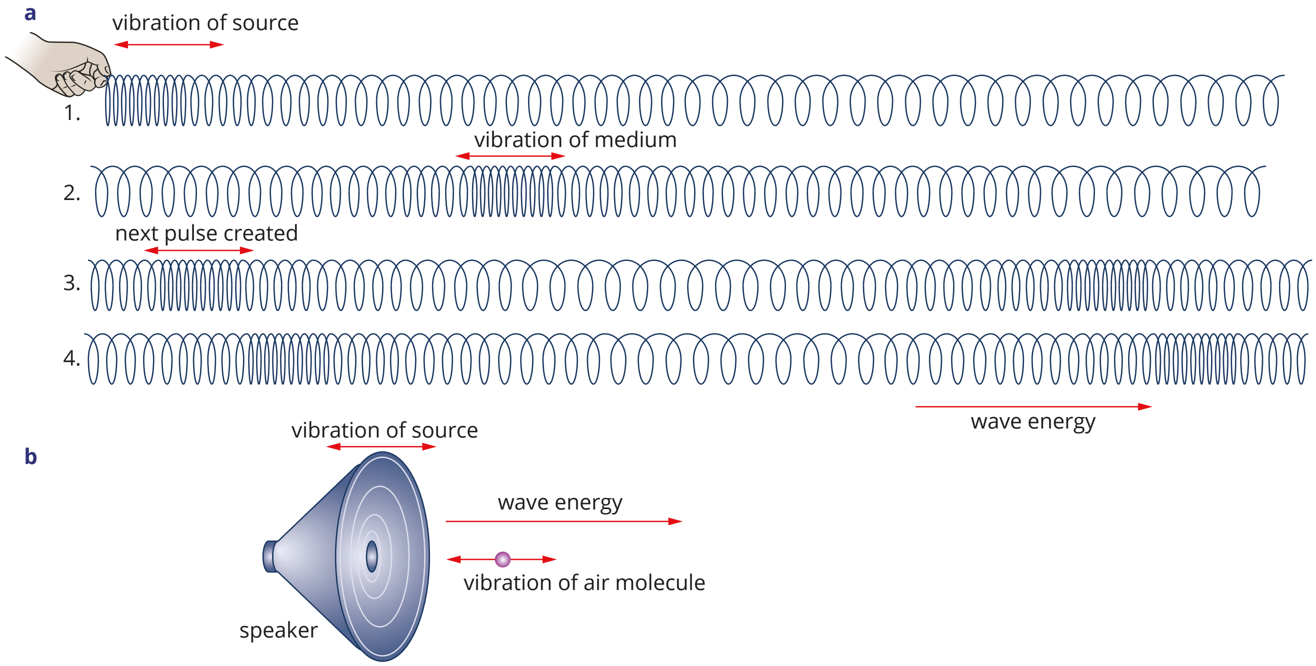


FIGURE 10.1.4 (a) When the direction of the vibrations of the medium and the direction of travel of the wave energy are parallel, a longitudinal wave is created. This can be demonstrated with a slinky. (b) Sound waves are longitudinal.

As you move your hand, a series of compressed and expanded areas form along the slinky. **Compressions** are those areas where the coils of the slinky come together. Expansions are regions where the coils are spread apart. Areas of expansion are called **rarefactions**. The compressions and rarefactions in a longitudinal wave correspond to the crests and troughs of a transverse wave.

An important example of a longitudinal wave is a sound wave, where small variations in **air pressure** carry the sound energy. As the cone of a loudspeaker vibrates (Fig. 10.1.4b), the layer of air next to it is alternately pushed away and drawn back creating a series of compressions (high pressure) and rarefactions (low pressure) in the air (demonstrated by the candle in Figure 10.1.5). This vibration is transmitted through the air as a sound wave, where the vibrations are in the direction of propagation. Like transverse waves, the individual particles undergo small vibrations back and forth about a mean position, while the wave itself can carry energy over very long distances. If the vibration was from a single point then the waves would tend to spread out spherically.

A longitudinal wave oscillates in the same direction as, or parallel to, the direction of the wave.

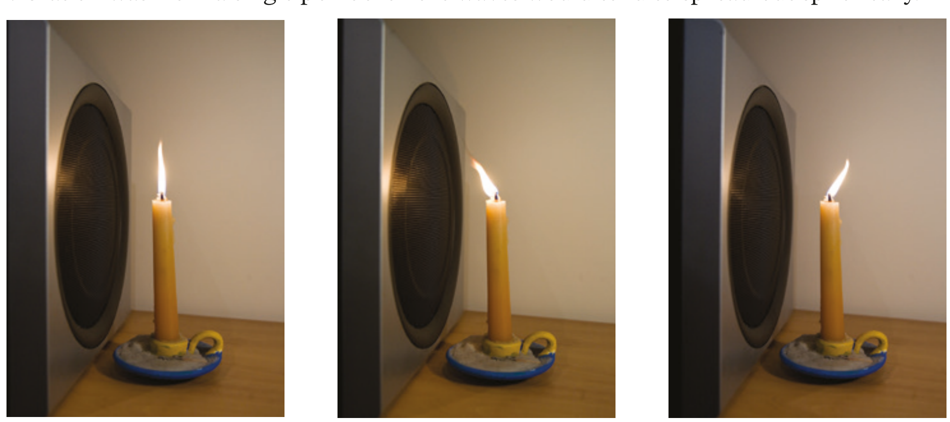


FIGURE 10.1.5 The motion of a flame in front of a loudspeaker is clear evidence of the continuous movement of air backwards and forwards as the loudspeaker creates a sound wave.

Measuring sound waves

It is important to remember that sound is a longitudinal wave. Sound waves are often measured by devices such as an oscilloscope, as shown in Figure 10.1.6. The sound waves are picked up by a microphone, then converted into an electrical signal that is then displayed as a transverse wave. Figure 10.1.7 shows the compressions (C) represented as a crest and the rarefactions (R) represented as a trough. For sound, the transverse wave display is only a convenient representation that makes it easier to determine properties such as amplitude, wavelength, frequency and period, discussed in the next section.

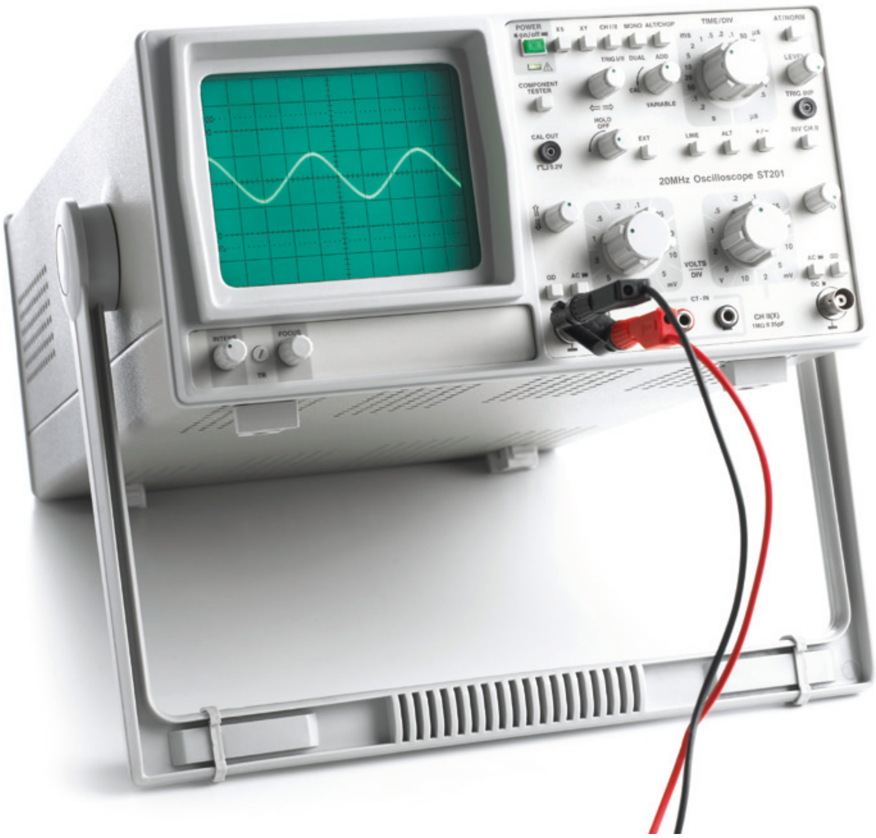


FIGURE 10.1.6 An oscilloscope measures the shape and parameters of electrical signals. Sound is converted to an electrical signal by a microphone and displayed on the oscilloscope.

PHYSICSFILE

Water waves

Water waves are often classified as transverse waves, but this is an approximation. In practical situations, transverse and longitudinal waves don't always occur in isolation. The breaking of waves on a beach produces complex wave forms which are a combination of transverse and longitudinal waves (Figure 10.1.8).

If you looked carefully at a cork bobbing about in gentle water waves you would notice that it doesn't move straight up and down but that it has a more elliptical motion. It moves up and down, and very slightly forwards and backwards as each wave passes. However, since this second aspect of the motion is so subtle, in most circumstances it is adequate to treat water waves as if they were purely transverse waves.

FIGURE 10.1.8 Even though this surfer rides forward on the wave, the water itself only moves in an elliptical motion as the wave passes.

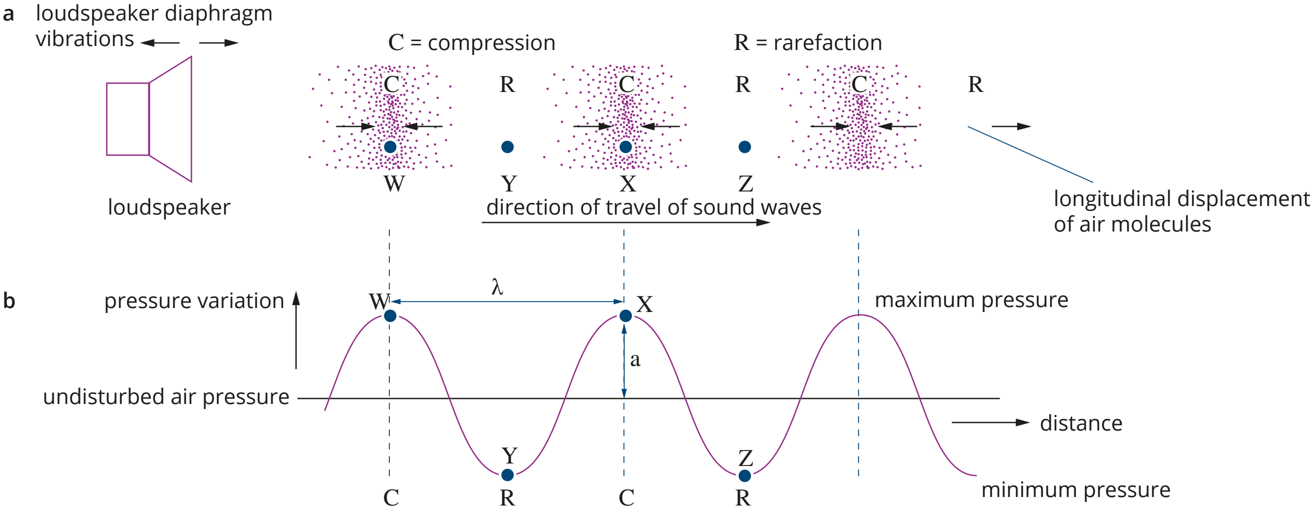


FIGURE 10.1.7 A longitudinal wave is compared to a transverse wave. The compressions in a longitudinal wave are equivalent to a crest or peak in a transverse wave, and the rarefactions are equivalent to a trough.

10.1 Review

SUMMARY

- Vibrating objects transfer energy through waves, travelling outwards from the source. Waves on water, on a string and sound waves in air are examples of mechanical waves.
- A wave may be a single pulse or it may be continuous or periodic (successive crests and troughs or compressions and rarefactions).
- A wave only transfers energy from one point to another. There is no net transfer of matter or material.
- The particles within the medium undergo small vibrations or oscillations about a mean position.
- Mechanical waves can be either transverse or longitudinal.
- In a transverse wave, the oscillations occur about a mean position and are perpendicular to the direction in which the wave energy is travelling. A wave in a string is an example of a transverse wave.
- In a longitudinal wave the particles or medium oscillate back and forth about a mean position. The oscillations are parallel to (along) the direction the wave energy is travelling. Sound is an example of a longitudinal wave.
- When longitudinal sound waves are converted to electrical signals by a microphone and displayed on an oscilloscope, the signal takes the shape of a transverse wave where the compressions are represented as a crest and the rarefactions as a trough.

KEY QUESTIONS

1 Describe the motion of particles within a medium as a mechanical wave passes through the medium.

2 State whether the following statements are true or false. For the false statements, rewrite them so they become true.

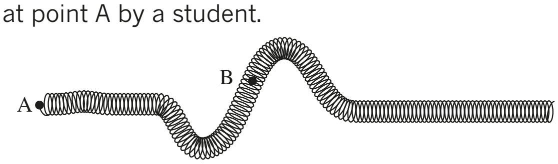
a Longitudinal waves occur when particles of the medium vibrate in the opposite direction to the direction of the wave.

b Transverse waves are created when the direction of vibration of the particles is at right angles to the direction of the wave.

c A longitudinal wave is able to travel through air.

d The vibrating string of a guitar is an example of a transverse wave.

3 The diagram below represents a slinky spring held at point A by a student.

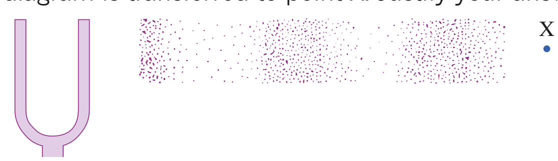


Draw an image of the pulse a short time after that shown in the diagram and determine the motion of point B. Will point B move upwards or downwards, or is it stationary?

4 Which of the following are examples of mechanical waves?

light, sound, ripples on a pond, vibrations in a rope

5 Describe how the energy from the tuning fork in the diagram is transferred to point X. Justify your answer.



6 The diagram below shows dots representing the average displacement of air particles at one moment in time as a sound wave travels to the right.

Describe how particles A and B have moved from their equally-spaced undisturbed positions to form the compression.

7 A mechanical wave may be described as transverse or longitudinal. In a *transverse* wave, how does the motion of the particles compare with the direction of travel of the wave?

8 Classify the waves described below as either longitudinal or transverse:

a sound waves

b a vibrating guitar string

c slinky moved with an upward pulse

d slinky pushed forwards and backwards.

10.2 Representing waves

In this section, you will explore how the displacement of particles within the wave can be represented using graphs. From these graphs, several key features of a wave can be identified:

- amplitude
- wavelength
- frequency
- period
- speed.

You will also learn how to do calculations using these features.

Transverse waves of different amplitudes and wavelengths can be seen in Figure 10.2.1.

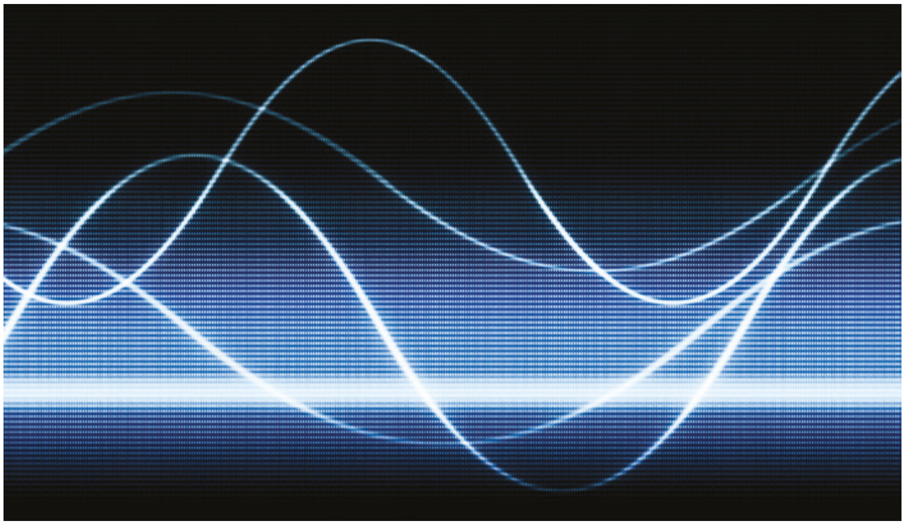


FIGURE 10.2.1 Waves can have different wavelengths, amplitudes, frequencies, periods and velocities, which can all be represented on a graph.

DISPLACEMENT–DISTANCE GRAPHS

The displacement–distance graph in Figure 10.2.2 shows the displacement of all particles along the length of a transverse wave at a particular point in time.

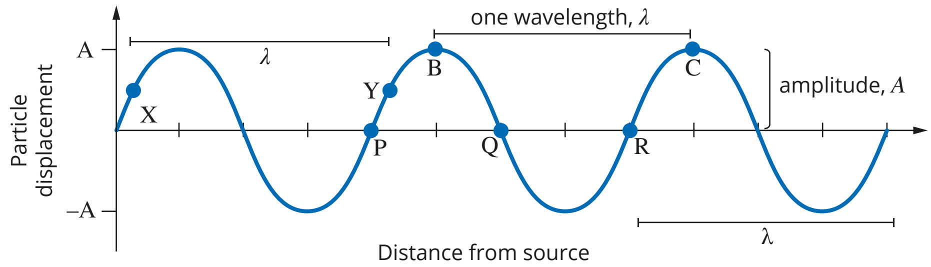


FIGURE 10.2.2 A sine wave representing the particle displacements along a wave over a distance.

Have a look back at Figure 10.1.2b on page 335 of a continuous wave in a slinky. This 'snapshot' in time shows the particles moving up and down **sinusoidally** about a central rest position. As a wave passes a given point, the particle at that point will go through a complete cycle before returning to its starting point. The wave spread along the length of the slinky has the shape of a sine or cosine function, which you will recognise from mathematics. A displacement–distance graph shows the position (displacement) of the particles along the slinky about a central position at a particular moment in time.

From a displacement–distance graph, the amplitude and wavelength of a wave are easily recognisable.

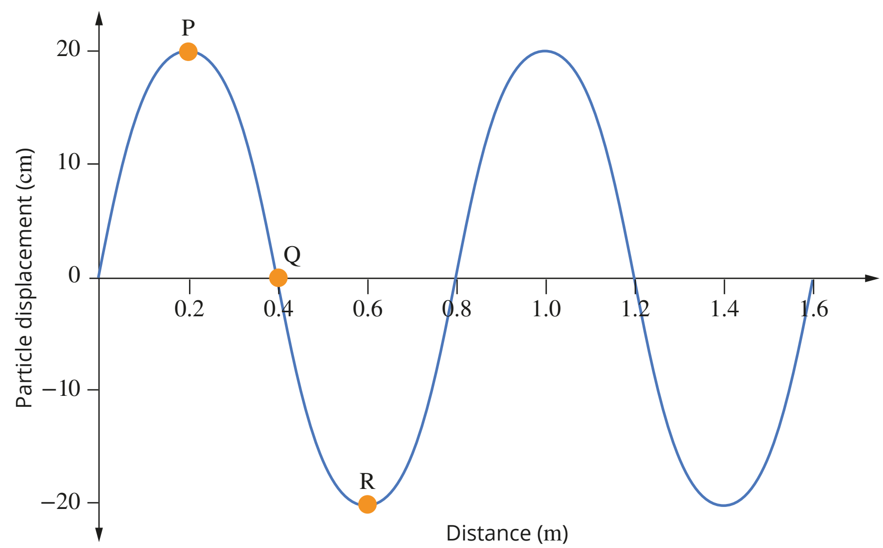
- The **amplitude** of a wave is the maximum displacement of a particle from the average or rest position, and is indicated on the vertical axis by A and –A. That is, the amplitude is distance from the middle of a wave to the top of a crest (A) or to the bottom of a trough (–A). The total distance a particle will move through in one cycle is twice the amplitude. The points B and C on the sinusoidal curve show the positions of the crests.
- The **wavelength** of a wave is the distance between any two successive points in phase (e.g. points B and C or X and Y in Figure 10.2.2). It is denoted by the Greek letter λ (lambda), and is measured in metres. Two particles on the wave are said to be in **phase** if they have the same displacements from the average position and are moving in the same direction. Points P and R in Figure 10.2.2 are two such particles that are in phase, as are points B and C and also X and Y. However, P and Q are not in phase as they are half a wavelength apart and are moving in opposite directions.

The **frequency**, *f*, is the number of complete cycles that pass a given point per second and is measured in hertz (Hz). By drawing a series of displacement–distance graphs at various times, you can see the motion of the wave. By comparing the changes in these graphs, the travelling speed and direction of the wave can be found, as well as the direction of motion of the vibrating particles.

Worked example 10.2.1

DISPLACEMENT–DISTANCE GRAPH

The displacement–distance graph below shows a snapshot of a transverse wave as it travels along a spring towards the right. Use the graph to determine the amplitude and the wavelength of this wave.

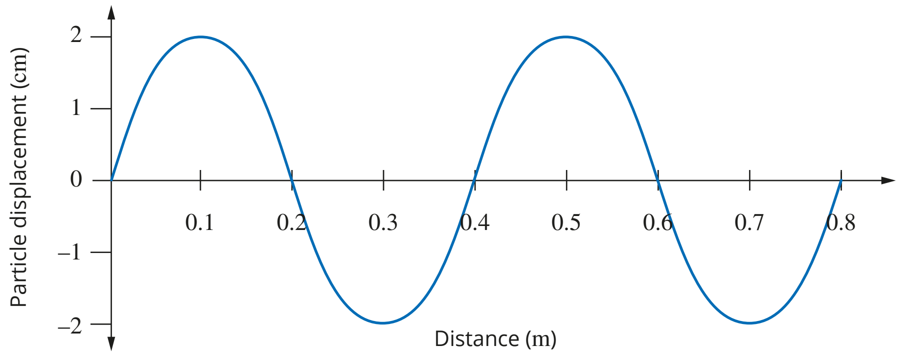


Thinking	Working
Amplitude on a displacement–distance graph is the distance from the average position (Q) to a crest (P) or a trough (R). Read the displacement of a crest or a trough from the vertical axis. Convert to SI units where necessary.	Amplitude = 20 cm = 0.2 m
Wavelength is the distance for one complete cycle. Any two consecutive points in phase and at the same position on the wave could be used.	The first cycle runs from the origin through P, Q, R to intersect the horizontal axis at 0.8 m. This intersection is the wavelength. Wavelength λ = 0.8 m

Worked example: Try yourself 10.2.1

DISPLACEMENT–DISTANCE GRAPH

The displacement–distance graph below shows a snapshot of a transverse wave as it travels along a spring towards the right. Use the graph to determine the wavelength and the amplitude of this wave.



DISPLACEMENT–TIME GRAPHS

A displacement–time graph such as the one shown in Figure 10.2.3 tracks the position of one point over time as the wave moves through that point.

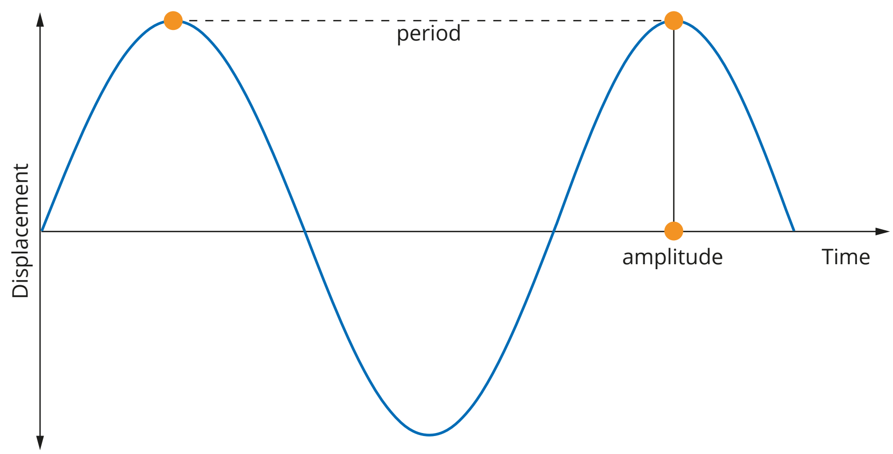


FIGURE 10.2.3 The graph of displacement versus time from the source of a transverse wave shows the movement of a *single point* on a wave over time as the wave passes through that point.

The displacement–time graph looks very similar to a displacement–distance graph of a transverse wave, so be careful to check the horizontal axis label.

Crests and troughs are shown the same way in both graphs. The amplitude is still the maximum displacement from the average or rest position of either a crest or a trough. But the distance between two successive points in phase in a displacement–time graph represents the **period** of the wave, T, measured in seconds.

The period is the time it takes for any point on the wave to go through one complete cycle (e.g. from crest to successive crest).

The period of a wave is inversely related to its frequency:

$$T = \frac{1}{f}$$

where T is the period of the wave (s)

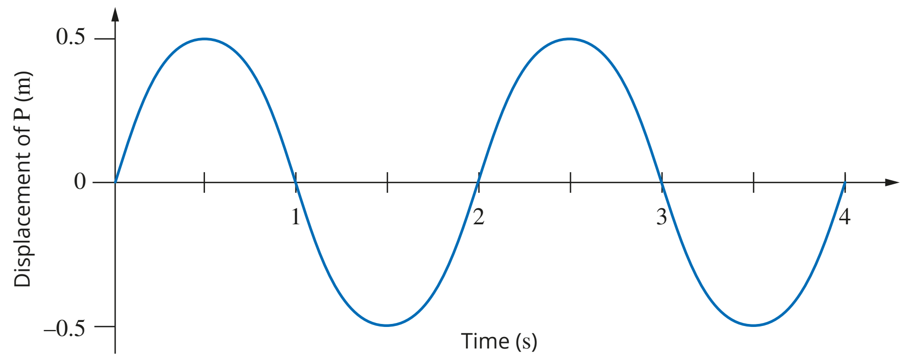
f is the frequency of the wave (Hz or s^{-1})

The amplitude and period of a wave can be determined directly from a displacement–time graph.

Worked example 10.2.2

DISPLACEMENT–TIME GRAPHS

The displacement–time graph below shows the motion of a single part of a rope (point P) as a wave passes by travelling to the right. Use the graph to find the amplitude, period and frequency of the wave.

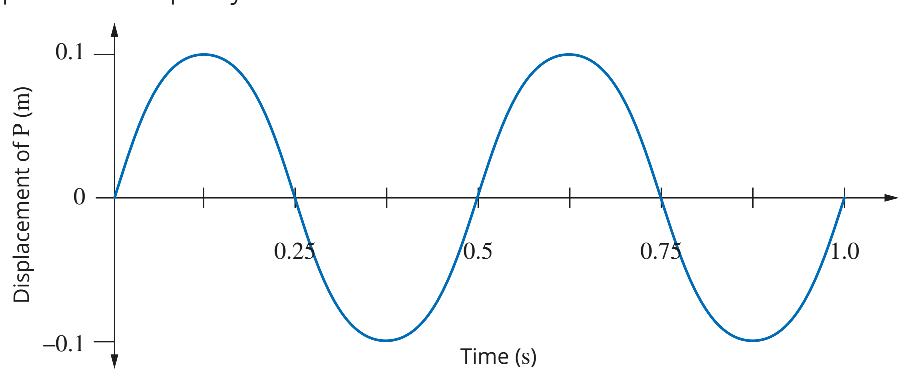


Thinking	Working
The amplitude on a displacement–time graph is the displacement from the average position to a crest or trough. Note the displacement of successive crests and/or troughs on the wave and carefully note units on the vertical axis.	Maximum displacement is 0.5 m Therefore amplitude = 0.5 m
Period is the time it takes to complete one cycle and can be identified on a displacement–time graph as the time between two successive points that are in phase. Identify two points on the graph at the same position in the wave cycle, e.g. the origin and $t = 2$ s. Confirm by checking two other points, e.g. two crests or two troughs.	Period, $T = 2$ s
Frequency can be calculated using $f = \frac{1}{T}$, measured in hertz (Hz).	$f = \frac{1}{T} = \frac{1}{2} = 0.5$ The frequency is 0.5 Hz.

Worked example: Try yourself 10.2.2

DISPLACEMENT–TIME GRAPHS

The displacement–time graph below shows the motion of a single part of a rope as a wave passes travelling to the right. Use the graph to find the amplitude, period and frequency of the wave.



Worked example 10.2.2 and Worked example: Try yourself 10.2.2 represent **travelling waves**. In these, every point on the rope would have a maximum displacement at some point as the wave travels down the rope, and similarly every point would have a minimum displacement at some point as the wave travels down the rope.

THE WAVE EQUATION

There is a relationship between the speed of a wave and the other significant wave characteristics of frequency, period and wavelength.

Think back to the study of motion in Chapter 7. Speed is given by:

$$v = \frac{\text{distance travelled}}{\text{time taken}} = \frac{d}{\Delta t}$$

This can be rewritten in terms of the distance of one wavelength (λ) in one period (T), which will be:

$$v = \frac{\lambda}{T}$$

and since

$$f = \frac{1}{T}$$

the relationship becomes:

$v = f\lambda$

where v is the speed ($m\,s^{-1}$)

f is the frequency (Hz)

λ is the wavelength (m)

This is known as the wave equation and applies to both longitudinal and transverse mechanical waves.

Worked example 10.2.3

THE WAVE EQUATION

A longitudinal wave has a wavelength of 2.00 m and a speed of 340 $m\,s^{-1}$. What is the frequency, f, of the wave?	
Thinking	**Working**
The wave equation states that $v = f\lambda$. Knowing both v and λ, the frequency, f, can be found. Rewrite the wave equation in terms of f.	$v = f\lambda$ $f = \frac{v}{\lambda}$
Substitute the known values and solve.	$f = \frac{v}{\lambda}$ $= \frac{340}{2.00}$ $= 170$ Hz

Worked example: Try yourself 10.2.3

THE WAVE EQUATION

A longitudinal wave has a wavelength of 3.00 m and a speed of 1484 $m\,s^{-1}$. What is the frequency, f, of the wave?

Worked example 10.2.4

THE WAVE EQUATION

A longitudinal wave has a wavelength of 2.00 m and a speed of 340 m s^{-1}. Calculate the period, T, of the wave.	
Thinking	**Working**
Rewrite the wave equation in terms of T.	$v = f\lambda$, and $f = \dfrac{1}{T}$ $v = \dfrac{1}{T} \times \lambda$ $= \dfrac{\lambda}{T}$ $T = \dfrac{\lambda}{v}$
Substitute the known values and solve.	$T = \dfrac{\lambda}{v}$ $= \dfrac{2.00}{340}$ $= 5.88 \times 10^{-3}$ s

Worked example: Try yourself 10.2.4

THE WAVE EQUATION

A longitudinal wave has a wavelength of 3.00 m and a speed of 1484 m s^{-1}. Calculate the period, T, of the wave.

THE DOPPLER EFFECT

The **Doppler effect** is a phenomenon of waves that is observed whenever there is relative movement between the source of the waves and an observer. Named after Austrian physicist Christian Doppler, who proposed it in 1842, the Doppler effect only affects the *apparent* frequency of the wave. The actual frequency of the wave does not change. A common experience of the Doppler effect is in listening to the sound of a siren from an emergency vehicle as it approaches and passes by.

Suppose a wave source, such as an ambulance siren, is stationary relative to an observer, as shown in the top section of Figure 10.2.4. The observer will receive and hear the disturbances (rarefactions and compressions in this example) at the same rate as the source creates them.

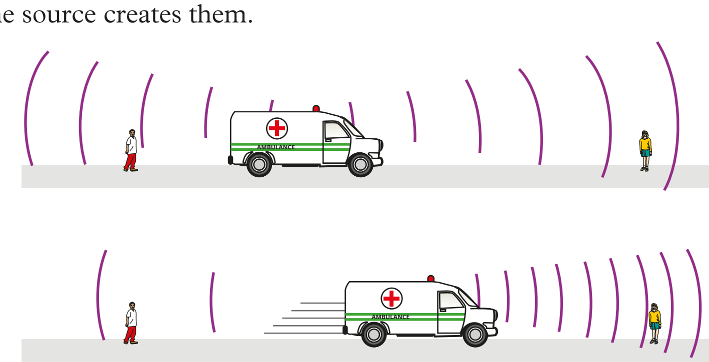


FIGURE 10.2.4 The Doppler effect. An object emitting a sound moving towards an observer (on the right) will emit sound waves closer together in its direction of travel and hence a higher frequency is heard by the observer. When the object is moving away from the observer (on the left), the sound waves are emitted further apart and hence a lower frequency is heard by the observer.

If, on the other hand, the wave source (the ambulance) were to travel towards the observer on the right in Figure 10.2.4, then each consecutive disturbance would originate from a position a little closer than the previous one. Hence each disturbance would have a little less distance to travel before reaching the observer than the one immediately before it. The effective wavelength of the wave fronts is less and therefore the frequency of arrival of the disturbances is higher than the originating frequency. The observer hears a higher frequency or pitch.

Alternatively, if the source (ambulance) is moving away from the observer (as shown on the left in Figure 10.2.4), then each consecutive disturbance will originate from a distance a little further away than the one immediately before and so has a greater distance to travel. The disturbances will arrive at the observer with a frequency that is less than the originating frequency, that is a lower pitch.

The net effect is that when the wave is moving towards an observer, the frequency of arrival of the wave will be higher than the frequency of the original source. When the wave is moving away from the observer, the frequency of arrival will be lower than the frequency of the original source. Therefore, as a result of the Doppler effect, a siren will appear to rise in frequency as the vehicle travels towards you and fall as it moves away.

For a mechanical wave, the total Doppler effect may result from the motion of the source, the motion of the observer, or the motion of the medium the wave travels through. For waves that don't require a medium, such as light, only the relative difference in speed between the observer and the source will contribute to the effect. For light, relativistic effects must be taken into account, as you will see in Year 12.

PHYSICSFILE

The sound of the Doppler effect

This Doppler effect behaviour can be easily modelled. You should be able to mimic the sound of a high-powered racing car like that in Figure 10.2.5 by making the sound 'neee…owwwww' with your voice. The 'neee' is the sound the racing car would make as it approached you—hence the high frequency. The 'owwwww' is the sound the racing car would make is it passed you and travelled away—hence the low frequency.

FIGURE 10.2.5 Australia's Daniel Ricciardo racing in Spain.

EXTENSION

Doppler calculations

This study only requires a qualitative understanding of the Doppler effect as a wave phenomenon. However, as the relative motion between the observer and source is the cause of the change in apparent frequency, then by knowing what the relative motion is, the apparent frequency can be calculated.

In classical physics, where both the speed of the source and the observer are lower than the speed of the waves in the medium, and the source and observer are approaching each other directly:

$$f = \left(\frac{v + v_0}{v - v_s}\right) f_0$$

where f is the apparent or observed frequency (Hz)

f_0 is the original frequency (Hz)

v is the speed of the waves in the medium (m s^{-1})

v_0 is the speed of the observer relative to the medium (m s^{-1}). v_0 is positive if the observer is moving towards the source and negative if the observer is moving away.

v_s is the speed of the source relative to the medium (m s^{-1}). v_s is positive if the source is moving towards the observer and negative if the source is moving away.

As an approximation, if the speeds of the source and observer are small relative to the speed of the wave, then the approximate observed frequency is:

$$f = \left(1 + \frac{\Delta v}{v}\right) f_0, \text{ where } \Delta v = v_0 - v_s$$

and the approximate apparent change in frequency is

$$\Delta f = \left(\frac{\Delta v}{v}\right) f_0, \text{ where } \Delta f = f - f_0$$

An interesting additional effect was predicted by British physicist Lord Rayleigh. He predicted that if the source is moving at double the speed of sound, a musical piece emitted by the source would be heard in correct time and frequency, but *backwards*. Try to establish whether his prediction is true mathematically using the formulas above.

10.2 Review

SUMMARY

- Waves can be represented by displacement–distance graphs and displacement–time graphs.
- From a displacement-distance graph you can directly determine the amplitude and wavelength of the wave.
- From a displacement–time graph, you can directly determine the amplitude and period of the wave.
- The period of a wave has an inverse relationship to its frequency, according to the relationship:

$$T = \frac{1}{f}$$

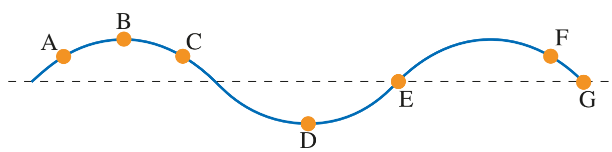
- The speed of a wave can be calculated using the wave equation:

$$v = f\lambda = \frac{\lambda}{T}$$

- The Doppler effect is a phenomenon that is observed whenever there is relative movement between the source of waves and an observer. It causes an apparent *increase* in frequency when the relative movement is *towards* the observer and an apparent *decrease* in frequency when the relative movement is *away* from the observer.
- For a mechanical wave, the total Doppler effect may result from the motion of the source, the motion of the observer, or the motion of the medium.

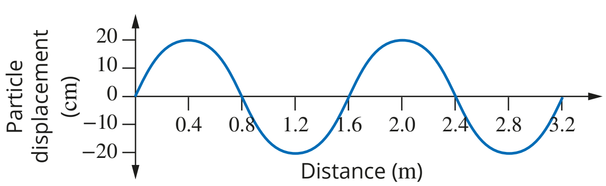
KEY QUESTIONS

1 Using the displacement–distance graph below, give the correct word or letters for the following:

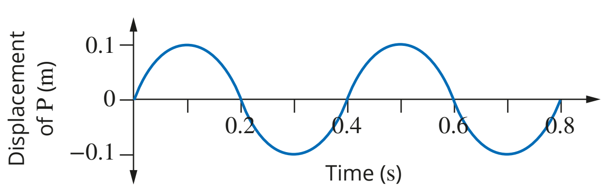


a two points on the wave that are in phase
b the name for the distance between these two points
c the two particles with maximum displacement from their rest position
d the term for this maximum displacement.

2 Use the graph below to determine the wavelength and the amplitude of this wave.



3 This is the displacement–time graph for a particle P.

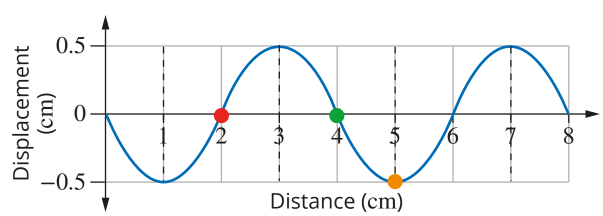


What is the:
a period of the wave
b frequency of the wave?

4 Five wavelengths of a wave pass a point each second. The amplitude is 0.3 m and the distance between successive crests of the waves is 1.3 m. What is the speed of the wave?

5 Which of the following is true and which is false? For the false statements, rewrite them to make them true.
a The frequency of a wave is inversely proportional to its wavelength.
b The period of a wave is inversely proportional to its wavelength.
c The amplitude of a wave is not related to its speed.
d Only the wavelength of a wave determines its speed.

6 Consider the displacement–distance graph below.



a State the wavelength and amplitude of the wave.
b If the wave moves through one wavelength in 2 s, what is the speed of the wave?
c If the wave is moving to the right, which of the coloured particles is moving down?

7 Calculate the period of a wave with frequency 2×10^5 Hz.

8 A police car, travelling at 100 km h^{-1} along a straight road, has its siren sounding. The police car is pursuing another car travelling in the same direction, also at 100 km h^{-1}. There is no wind at the time. Would an observer in the car being pursued hear the siren from the police car at a higher, lower or the same frequency as it emits? Explain your answer.

9 An ambulance sounding its siren in still air moves towards you, then passes you and continues to move away in a straight path. How would the siren sound to you?

10.3 Wave behaviours—reflection, refraction and diffraction

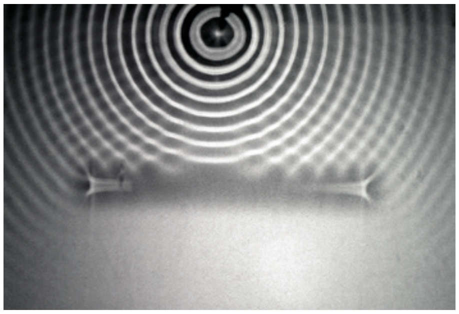


FIGURE 10.3.1 The reflection of spherical waves in a ripple tank when meeting a solid surface; in this case a barrier is positioned below the source of the circular waves.

Mechanical waves transfer energy through a medium, but there will be times when that medium physically ends, such as when a water meets the edge of a pool or air meets a wall. A change in the physical characteristics in the same medium, such as density and temperature, can act like a change in medium. When the medium ends, or changes, the wave doesn't just stop. Instead, the energy that the wave is carrying will undergo three processes:

- some energy will be *reflected* (Figure 10.3.1)
- some energy will be *absorbed* by the new medium
- some energy will be *transmitted*.

The degree to which each occurs will depend upon the differences in properties between the original and new media. Energy transfers and transformations can be understood by investigating wave behaviours.

REFLECTION

When a transverse wave pulse reaches a hard surface, such as the fixed end of a rope, the wave is bounced back or **reflected**.

Reflection characteristics

When the end of the rope is fixed, the reflected pulse is inverted (Figure 10.3.2). So, for example, a wave crest would be reflected as a trough.

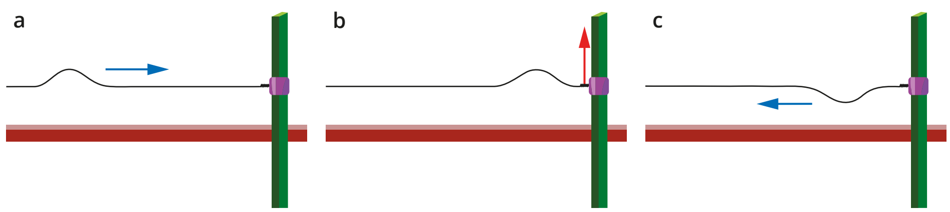


FIGURE 10.3.2 (a) A wave pulse moves along a string to the right and approaches a fixed post. (b) On reaching the end, the string exerts an upwards force on the fixed post. Due to Newton's third law, the fixed post exerts an equal and opposite force on the string which (c) inverts the wave pulse and sends its reflection back to the left on the bottom side of the string. There is a phase reversal on reflection from a fixed end.

This inversion can also be referred to as a 180° change of phase or, expressed in terms of the wavelength, λ, a phase shift of $\frac{\lambda}{2}$.

When a wave pulse hits the end of a rope that is free to move (known as a free boundary), the pulse returns with no change of phase (Figure 10.3.3). That is, the reflected pulse is the same as the incident pulse. A crest is reflected as a crest and a trough is reflected as a trough.

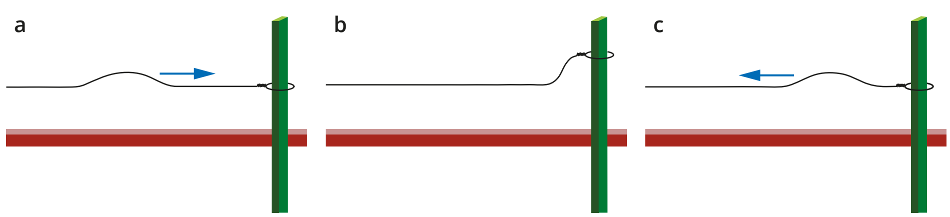


FIGURE 10.3.3 (a) A wave pulse moves along a string to the right and approaches a free end at the post. (b) On reaching the post the free end of the string is free to slide up the post. (c) No inversion happens and the wave pulse is reflected back to the left on the same side of the string, i.e. there is no phase reversal on reflection from a free end.

When the transverse wave pulse is reflected, the amplitude of the reflected wave isn't quite the same as the original. Part of the energy of the wave is **absorbed** by the post, where some will be transformed into heat energy and some will continue to travel through the post. You can see this more clearly by connecting a heavier rope to a lighter rope. The change in density has the same effect as a change in medium (Figure 10.3.4).

When a transverse wave pulse is sent down the rope from the light rope to the heavier rope, part of the wave pulse will be reflected and part of it will be **transmitted** to the heavier rope. As the second rope is heavier, a smaller proportion of the wave is transmitted into it and a larger proportion of the wave is reflected back.

This is just the same as a wave pulse striking a wall. The more rigid and/or dense the wall, the more the wave energy will be reflected and the less it will be absorbed – but there will always be some energy that is absorbed by or transferred to the second medium. This explains why sound can travel through walls.

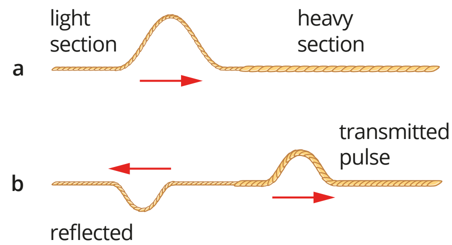


FIGURE 10.3.4 (a) A wave pulse travels along a light rope towards a heavier rope. (b) On reaching a change in density the wave pulse will be partly reflected and partly transmitted. This is analogous to a change in medium.

Reflected wave fronts

Two- and three-dimensional waves, such as water waves, travel as **wave fronts**. When drawing wave fronts (Figure 10.3.5), it is common to show the crests of the waves. When close to the source, wave fronts can show considerable curvature (Figure 10.3.5a) or may even be spherical when generated in three dimensions. Where a wave has travelled a long distance from its source, the wave front is nearly straight and is called a **plane wave**. A plane wave is shown in Figure 10.3.5b. Plane waves can also be generated by a long, flat source such as those often used in a ripple tank.

The direction of motion of any wave front can be represented by a line drawn perpendicular to the wave front and in the direction the wave is as shown by the blue arrows in Figure 10.3.5. This is called a **ray**. Rays can be used to study or illustrate the properties of two- and three-dimensional waves without the need to draw individual wave fronts.

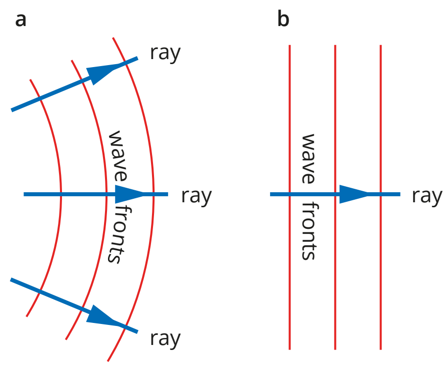


FIGURE 10.3.5 Rays can be used to illustrate the direction of motion of a wave. They are drawn perpendicular to the wave front of a two- or three-dimensional wave and in the direction of travel of the wave; (a) illustrates rays for circular waves near a point source while (b) shows a ray for plane waves.

By using rays to illustrate the path of a wave front reflecting from a surface, it can be shown that for a two- or three-dimensional wave, the angle from the **normal** at which the wave strikes a surface will equal the angle from the normal to the reflected wave. The normal is an imaginary line at 90°, i.e. perpendicular, to the surface.

These angles of the incident and reflected waves from the normal are labelled θ_i and θ_r, respectively, in Figure 10.3.6.

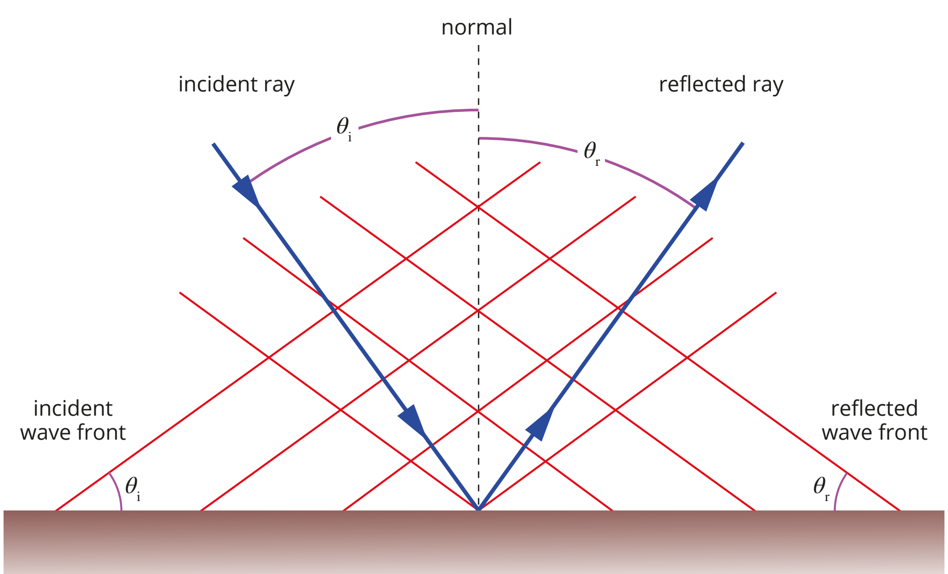


FIGURE 10.3.6 The law of reflection. The angle between the incident ray and the normal (θ_i) is the same as the angle between the normal and the reflected ray (θ_r).

This is referred to as the law of reflection.

The law of reflection states that the **angle of reflection** (θ_r), measured from the normal, equals the **angle of incidence** (θ_i) measured from the normal

$\theta_i = \theta_r$.

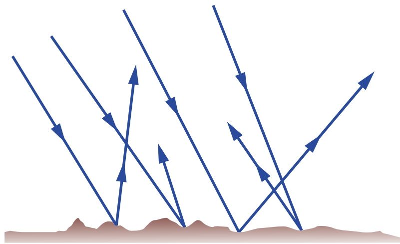


FIGURE 10.3.7 Reflection from an irregular surface. Each incident ray may be reflected in a different direction, depending upon how rough or irregular the reflecting surface is. The resulting wave will be diffuse (spread out).

The law of reflection is true for any surface whether it is straight, curved or irregular. For all surfaces, including curved or irregular surfaces, the normal is drawn perpendicular to the surface at the point of contact of the incident ray or rays.

When wave fronts meet an irregular, rough surface, the resulting reflection can be spread over a broad area. This is because each point on the surface may reflect the portion of the wave front reaching it in a different direction, as seen in Figure 10.3.7. This is referred to as **diffuse** reflection.

Echoes and reverberation

Echoes provide the most obvious evidence that sound waves are reflected. Like all waves, sound can be reflected when it strikes an obstacle.

When a sound wave that is reflected at a wall reaches us in a time of about 0.1 seconds or more you can hear an echo, and it is heard as a separate sound to the original sound. In smaller rooms, the waves that are reflected several times at different walls overlap and produce **reverberation**, which often sounds like a longer sound. Echoes typically occur in large rooms or large caverns with bare walls and hard surfaces. You may have noticed this in a vacant house. Simple ways to reduce echoes include using acoustic foam sheets, rugs on the floor, textiles on the wall to absorb the sound and objects in the room to diffuse the reflections of the sound waves. This is a form of acoustic engineering.

PHYSICS IN ACTION

Sonic Depth Finder

An echo-sounding device, called a sonar, makes it possible to measure the depth of the sea (Figure 10.3.8). A sound wave is emitted from the device and is reflected by the sea bed. The device measures the time taken for the echo to return to the ship, and knowing the speed of sound in water determines depth.

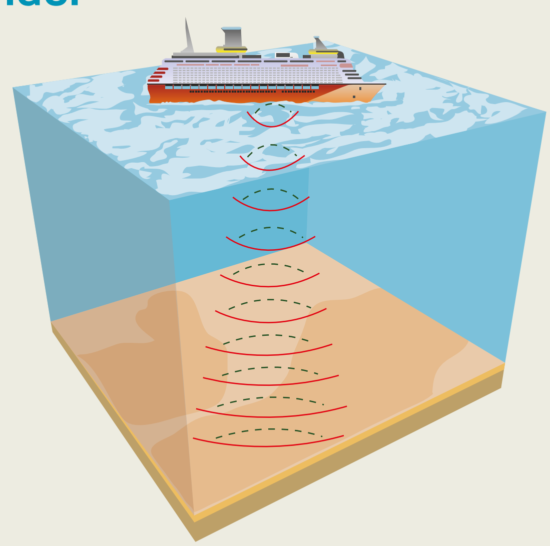


FIGURE 10.3.8 The ship sends a sound wave into the water below the ship. By measuring the time between the emitted sound (red waves) and its reflection (dashed lines), the ship's sonar can determine the depth of the water under the ship.

REFRACTION

As discussed previously, when a wave hits a boundary between two media, some of the wave's energy is reflected, some is transmitted and some is absorbed by the new medium. The velocity of the wave absorbed by the second medium will be affected by the properties of the new medium, which will also affect the wavelength of the wave in the new medium.

When a wave crosses the boundary from one medium to another, and where there is a change in velocity and wavelength, the transmitted wave travels in a different direction to the original wave. This is referred to as **refraction**, and is the same phenomenon observed for light when a spoon appears broken at the interface between air and water in a glass. In Figure 10.3.9 the wave is travelling faster in the first medium and slower in the second medium.

Wave fronts refract because the part of the wave front that reaches the boundary first slows down (or speeds up), while the remainder of the wave front still in the original medium continues at its original speed.

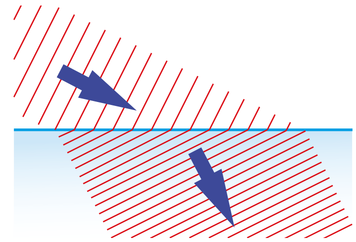


FIGURE 10.3.9 Refraction of wave fronts passing the boundary from one medium to another occurs when the speed of the wave differs between the two media. In this figure, the wave is travelling faster in the first medium and slower in the second medium.

The speed of sound

The speed of sound can be quite different in various media and depends on two main parameters: stiffness or **elasticity**, and inertia or density. The elasticity of a medium is a measure of the tendency of a medium to return to its original shape after it has been bent, compressed or stretched. When sound propagates through a medium in a series of compressions and rarefactions, it will travel faster in a more elastic medium. Therefore, a solid is significantly more elastic than a liquid, and a liquid is more elastic than a gas. If there is a change in state as a sound wave travels from one medium to the next, there will be a significant effect on the speed.

The density also has an effect on the speed. The more dense the medium is, the slower the speed. This is because the greater mass per unit volume makes it more difficult for a given force to move the medium, so the vibrations take more time.

The speed of sound depends on the medium's stiffness or elasticity, called Young's modulus, E, and its density, ρ. It is given by the expression:

$$v = \sqrt{\frac{E}{\rho}}$$

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

E is Young's modulus (N m^{-2})

ρ is the density of the medium (kg m^{-3}).

The speed of sound for a selection of substances is given in Table 10.3.1.

TABLE 10.3.1 The speed of sound for a selection of substances.

State of matter	Substance	Speed of sound
Solid	diamond	$12000\ \text{m s}^{-1}$
	copper	$3560\ \text{m s}^{-1}$
Liquid	water	$1493\ \text{m s}^{-1}$
	mercury	$1450\ \text{m s}^{-1}$
	kerosene	$1324\ \text{m s}^{-1}$
Gas	air (20°C)	$343\ \text{m s}^{-1}$
	helium (0°C)	$972\ \text{m s}^{-1}$
	hydrogen (0°C)	$1286\ \text{m s}^{-1}$

Temperature also affects the speed of sound.

Increasing the temperature of a medium reduces its density without changing its elasticity, so it increases the speed that sound can travel through it.

Conversely decreasing the temperature of a medium increases its density, therefore decreasing the speed that sound can travel through it. Refraction can therefore occur at the interface between a hot and a cold medium.

Refraction of sound

Figure 10.3.10 shows examples of refraction for sound waves. The normal line, in green, is drawn at 90° to the boundary between the two media and is shown by the dotted line. The refraction of the wave is measured by comparing the **angle of incidence**, i to the **angle of refraction**, r. The change in direction of the wave is described as being towards the normal or away from the normal when compared to the angle of incidence.

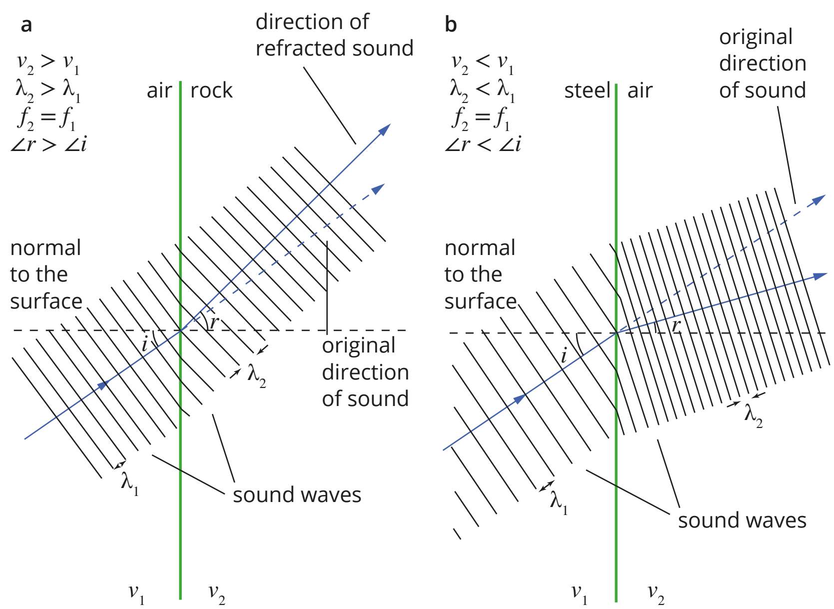


FIGURE 10.3.10 (a) The speed of the wave front in the second medium (the rock) is faster than in the first medium (air) and the wave front bends away from the normal. (b) The speed of the wave front in the second medium (air) is slower than in the first medium (steel) and the wave front bends towards the normal.

In Figure 10.3.10a the sound wave is travelling from air into rock. Since rock is a solid, the speed of sound in the rock, v_2, is greater than the speed of sound in air, v_1. The frequency of the waves ($f_1 = f_2$) and hence the time between wave fronts, T, remains constant in both media. Recalling the general expression, distance = velocity × time, the distance between the wave fronts is given by:

$$\lambda = vT$$

where $T = \frac{1}{f}$. Therefore, since the velocity of the wave in the rock is greater ($v_2 > v_1$), the spacing between the wave fronts (the wavelength) will increase ($\lambda_2 > \lambda_1$). This will cause the angle of refraction, r, to increase away from the normal ($\angle r > \angle i$).

In Figure 10.3.10b the sound wave is travelling from steel into air. In this case, the speed of sound in air is less than in steel ($v_2 < v_1$). Therefore, the spacing between the wave fronts will decrease ($\lambda_2 < \lambda_1$) and the angle of refraction, r, will decrease towards the normal ($\angle r < \angle i$).

The relationship can be stated mathematically using Snell's law, named after Dutch astronomer and mathematician Willebrord Snell:

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

To summarise:

When a wave front *speeds* up at the boundary between two media, the *wavelength increases* and the wave front will *bend away* from the normal.

When a wave front *slows down* at the boundary between two media, the *wavelength decreases* and the wave front will *bend towards* the normal.

The frequency of the wave is unaffected.

A wave front reaching a boundary perpendicular to the normal will not be affected and will continue in the same direction, regardless of whether it is speeding up or slowing down.

Worked example 10.3.1

REFRACTION

A 500 Hz sound from a rock concert travels from cold air where the speed of sound is 330 m s^{-1} into warmer air where the speed of sound is 355 m s^{-1} and strikes the boundary between the two regions of air with an angle of incidence of 30°.

a Find the wavelength in each region of air.	
Thinking	**Working**
The wavelength can be determined from $v = f\lambda$. Rearrange to make λ the subject: $\lambda = \frac{v}{f}$ where $f = 500$ Hz $v_{\text{cold air}} = 330\,\text{m s}^{-1}$ $v_{\text{warm air}} = 355\,\text{m s}^{-1}$	$\lambda_{\text{cold air}} = \frac{330}{500} = 0.60\,\text{m}$ $\lambda_{\text{warm air}} = \frac{355}{500} = 0.71\,\text{m}$

b Explain what will happen to the angle of refraction, and why.	
The speed of sound in warm air is higher than in cold air due to the lower air density in the warmer air.	Since the velocity in warm air is higher than in cold air, the angle will refract away from the normal.

c Calculate the angle of refraction.	
The angle of refraction can be calculated using Snell's law. $\frac{\sin r}{\sin i} = \frac{v_{\text{warm air}}}{v_{\text{cold air}}}$	$\sin r = \frac{v_{\text{warm air}}}{v_{\text{cold air}}} \sin i$ $\sin r = \frac{355}{340} \sin 30°$ $= 0.522$ $r = 31.5°$

Worked example: Try yourself 10.3.1

REFRACTION

Whales typically emit sounds between 10 and 40 Hz (humans can usually hear down to 20 Hz). If a whale emits a 20Hz sound in water towards the surface at an angle of 40° to the normal, the refracted wave emerges from the water into air. The speed of sound in air is 343 m s^{-1} and the speed of sound in water is 1484 m s^{-1}.

a Determine the wavelength in water and in air.

b Explain what will happen to the refracted wave, and why.

c Determine the angle of refraction.

It should be noted that whenever refraction occurs, a proportion of the wave is also reflected as shown by the ray diagram in Figure 10.3.11. For example, when light is incident on a window most of the light penetrates the window and is refracted into the glass, however however approximately 4% of the light is reflected.

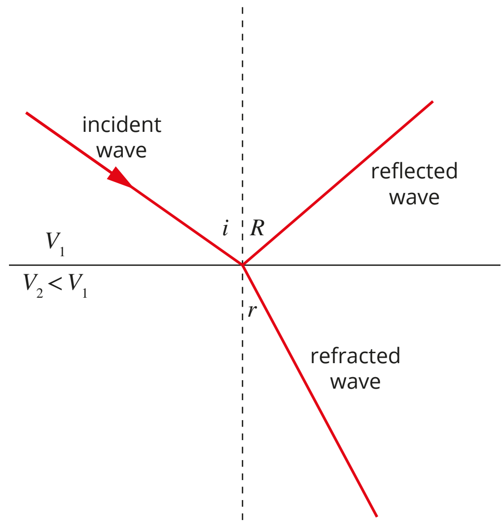


FIGURE 10.3.11 When a wave is incident on a new medium some of the wave is refracted but a proportion is also reflected.

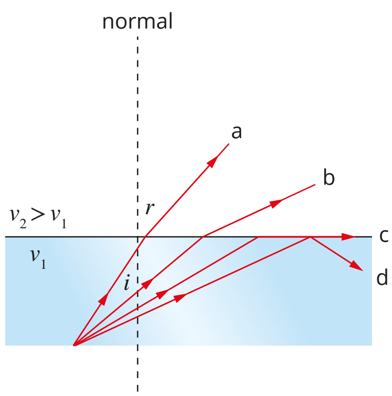


FIGURE 10.3.12 (a) and (b) show the refracted angle, *r*, getting increasingly larger as the incident angle, *i*, increases. (c) shows the critical angle where total internal reflection occurs.
(d) At an incident angle greater than the critical angle, reflection occurs.

Total internal reflection

As discussed in this section, when sound enters a medium where its speed is greater, the angle of refraction increases. It follows that as the incident angle gets increasingly larger then so will the reflected angle, as shown by the ray diagram in Figure 10.3.12a and b. At a particular incident angle known as the **critical angle**, θ_c, the refracted angle, r, will be 90°; that is, it will lie exactly along the interface between the two media, as shown in Figure 10.3.12c. This is known as **total internal reflection**. At incident angles greater than the critical angle, the ray is completely reflected (Figure 10.3.12d).

This can be shown mathematically through Snell's law. If the angle $r = 90°$, then $\sin r = 1$ and $i = \theta_c$ and Snell's law

$$\frac{\sin r}{\sin i} = \frac{v_2}{v_1} \text{ becomes } \sin\theta_c = \frac{v_1}{v_2}$$

Worked example 10.3.2

TOTAL INTERNAL REFLECTION

At what minimum incident angle would sound need to strike water from air if it is to reflect completely? The speed of sound in air is $344\,\text{m s}^{-1}$ and in water is $1500\,\text{m s}^{-1}$.

Thinking	Working
The critical angle needs to be calculated using Snell's law $\frac{\sin r}{\sin i} = \frac{v_2}{v_1}$ Set $\sin r = 1$ $v_1 = 344\,\text{m s}^{-1}$ $v_2 = 1500\,\text{m s}^{-1}$	$\sin\theta_c = \frac{v_1}{v_2}$ $\sin\theta_c = \frac{344}{1500} = 0.229$ $\theta_c = 13.3°$

Worked example: Try yourself 10.3.2

TOTAL INTERNAL REFLECTION

Sound is travelling through air and hits a steel wall. At what angle is the sound totally reflected? The speed of sound in steel is $5000\,m\,s^{-1}$ and the speed of sound in air is $340\,m\,s^{-1}$.

The critical angle is quite small if there is a large difference between the speed of sound in the first medium and the speed of sound in the second medium. In the example of sound going through air and striking water, the speed of sound in the second medium was five times that of the first, giving a critical angle of only 13.3°. Therefore, if the incident angle is greater than 13.3° the sound would be completely reflected.

In Worked example: Try yourself 10.3.2, where the second medium is a solid wall of steel and the speed of sound in steel is $5000\,m\,s^{-1}$, almost 15 times that of air, the critical angle is only about 4°. This explains why liquids and solids are good reflectors of sound—the sound will only penetrate at a very narrow range of incident angles. Thus, solid walls are a simple example of acoustic engineering where they are commonly used as sound insulators (Figure 10.3.13).

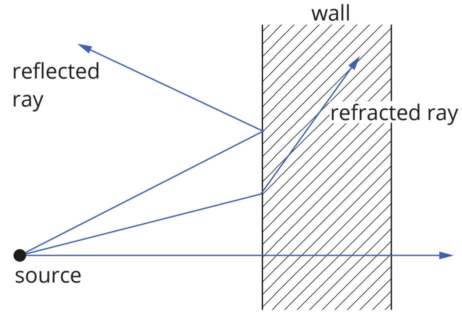


FIGURE 10.3.13 A wall acts as a good sound insulator, because sound transmitted through air towards a wall will only refract through a very narrow range of incident angles. At all incident angles greater than this critical angle sound will be reflected.

SEISMIC WAVES

When an earthquake occurs, it creates **seismic waves**. A seismic wave can take many forms. The main body wave of an earthquake travels through the Earth and consists of both primary longitudinal pressure waves (*P waves* for primary or pressure) and secondary transverse shear waves (*S waves* for shear or secondary). While both longitudinal and transverse waves can travel through a solid, only longitudinal waves can propagate through a fluid.

In fact, geophysicists have noted that both types of seismic waves, P and S, are detected close to the earthquake epicentre. However, while longitudinal P waves are detected at most places on the other side of the globe, the transverse S waves are not. There are large 'shadows' where transverse S waves are unable to propagate. As transverse waves cannot travel through a liquid this implies that the Earth's outer core is liquid. This is also supported by the fact that the molten core refracts the transverse P waves. Figure 10.3.14 shows where in the Earth's surface P and S waves are detected following an earthquake.

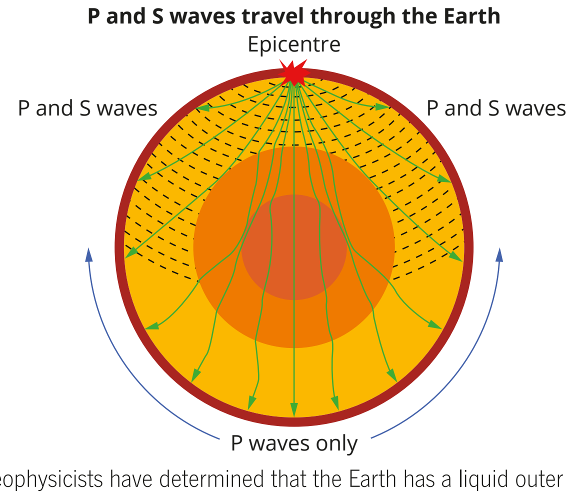


FIGURE 10.3.14 Geophysicists have determined that the Earth has a liquid outer core by observing the behaviour of seismic waves travelling through the Earth. Transverse S waves cannot travel through a liquid and are blocked by the liquid outer core while longitudinal P waves can be observed travelling through all of the Earth's interior.

DIFFRACTION OF SOUND

If you call someone from around the corner of a building they can still hear you even though they can't see you. This well-known ability of sound to travel around corners provides further evidence that sound is wave-like in nature. Reflection alone cannot account for all the indirect sounds. In addition, higher frequency sounds can be heard more clearly if the listener is directly in front of the source, whereas low frequency sounds can be heard quite clearly from a wide range of angles. For example, if you drive past someone who has their car stereo volume up very loud, the dominant sounds you will hear are the bass or low frequency sounds. This phenomenon is known as **diffraction.**

Diffraction effects in water as a plane wave passes through a narrow aperture are shown in Figure 10.3.15a and b. The interference effects shown in Figure 10.3.15b are discussed in Section 10.4.

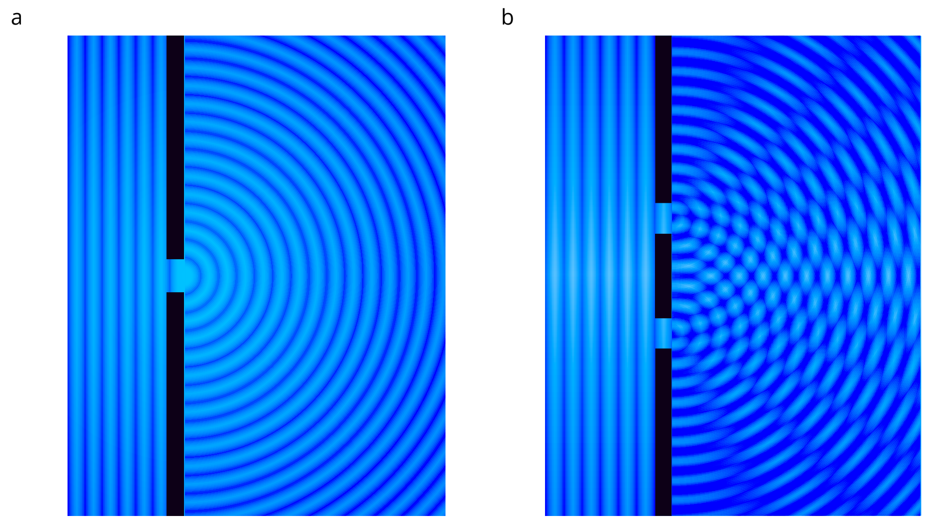


FIGURE 10.3.15 (a) A plane wave undergoes diffraction after passing through an aperture. (b) Interference effects occur due to diffraction from both of the double slits.

Diffraction is the bending of waves as they pass the edge (or edges) of an obstacle or pass through an aperture.

Diffraction and wavelength

How much a sound wave spreads out on passing an obstacle or through an aperture will depend upon the size of the wavelength of the sound, λ, in relation to the width of the aperture or obstacle, *d*. If the aperture width is similar to or much less than the wavelength, diffraction effects become significant, as shown in Figure 10.3.16a and b where there is significant spreading of the waves. However, if the aperture width, or the barrier width, is much larger than the wavelength then the amount of diffraction is very limited, as shown in Figure 10 3.16c and d where there is minimal disturbance to the wave.

Significant diffraction will occur when the wavelength is of the same order of magnitude or larger than the width of the obstacle or aperture.

Diffraction is therefore more significant at longer wavelengths and is the reason why it is difficult to pinpoint the exact source of long wavelength (low frequency) sound. The wavelengths of sound within normal hearing range are between about 2 cm and 20 m. A typical human voice has a wavelength of around 1 m, so voices diffract easily through doorways and around large obstacles.

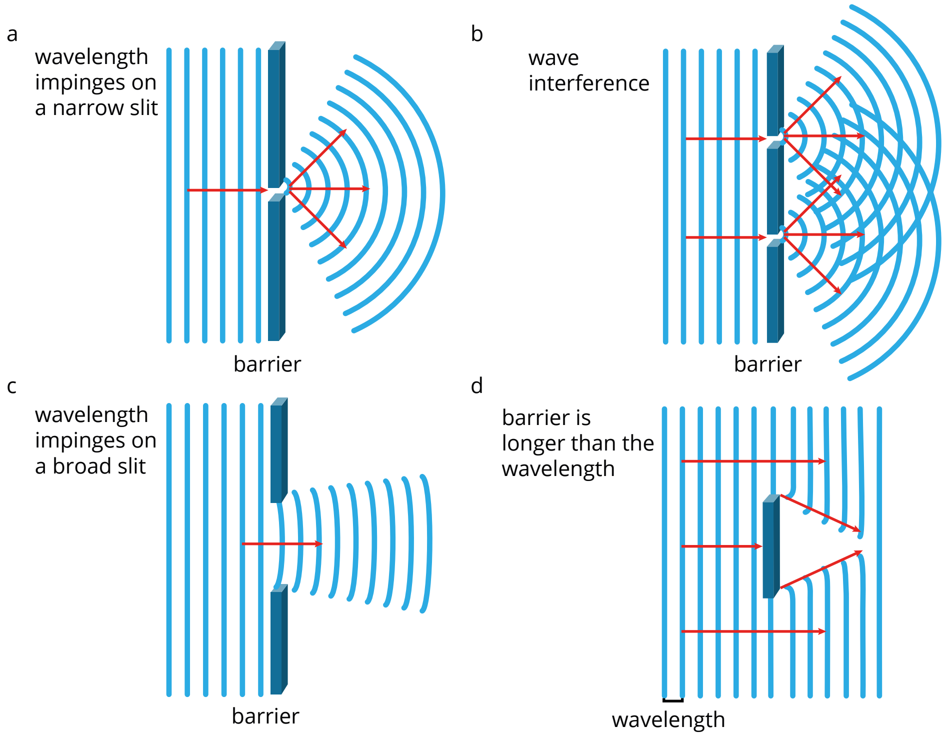


FIGURE 10.3.16 The amount of diffraction depends on the size of the gap or aperture and the wavelength. When the aperture is much smaller than the wavelength, significant diffraction, or spreading out of the waves, will occur ((a) and (b)). A large aperture (c) or a wide barrier (d) relative to the wavelength will cause little diffraction (d). Interference effects occur due to the superposition of two diffracted waves (b).

Diffraction and applications

As discussed previously, the wave equation is $v = f\lambda$, which gives f inversely proportional to λ. Therefore, high frequency (short wavelength) sounds are diffracted less, they are more directional, and it is easier to hear them from a particular direction. Ultrasound (with frequencies greater than 20 000 Hz) are used for sonar and ultrasonic motion detectors because its short wavelength means that diffraction is very limited. The sonar beam tends to travel to and from an object with only a small degree of spread. Figure 10.3.17 shows a dolphin locating its prey using sonar.

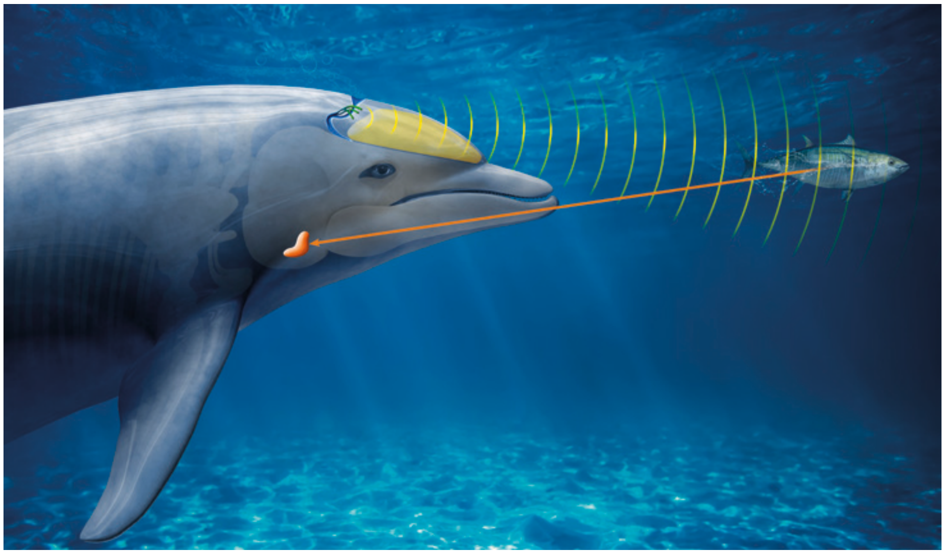


FIGURE 10.3.17 A dolphin locates its prey using echolocation or sonar.

Lower frequency sounds, with larger wavelengths, are significantly diffracted and therefore will readily fill a room. Therefore when setting up a stereo surround sound system, one correctly placed sub-woofer (low-frequency speaker) is usually adequate, whereas the mid- and high-frequency speakers need to be placed around the listener. This is not the only consideration when setting up a sound system as will be discussed later in this chapter.

Worked example 10.3.3

DIFFRACTION

When high frequency sound waves, e.g. 9000 Hz strike an obstacle such as a person's head, they leave a distinct sound shadow in which little of the sound can be heard. If one ear is closer to the sound source than the other, these higher frequency sounds will be heard as louder by the ear closer to the source.	
a Assuming the speed of sound is 340 m s^{-1} calculate the wavelength of the 9000 Hz sound.	
Thinking	**Working**
The formula is $v = f\lambda$. Therefore $\lambda = \frac{v}{f}$.	$\lambda = \frac{340}{9000}$ $= 3.78 \times 10^{-2}$ m

b Use this calculation to explain why this frequency leaves a sound shadow.	
Compare the wavelength of sound to the distance between a person's ears (approximately 20 cm).	$\lambda = 3.78$ cm is much less than 20 cm. Therefore, diffraction effects are minimal.

Worked example: Try yourself 10.3.3

DIFFRACTION

In ultrasound imaging the speed of sound is 1540 m s^{-1}. The resolution of an image depends on the wavelength of the sound—a smaller wavelength (higher frequency), enables more detail to be seen with less effect of diffraction. High-frequency sound (5 to 10 MHz) can resolve more detail but has limited penetration depth whereas low-frequency sound (2 to 5 MHz) can penetrate to deeper structures but has lower resolution.

a If the human heart is 10 cm across, what frequency is needed to have at least 300 wavelengths across the image?

b If the frequency were significantly lower than your calculated amount, what would happen to the image? Explain why.

PHYSICSFILE

Hearing Loss

As a person gets older, or if they have been continually exposed to excessive noise, such as overloud earphones, they tend to lose the ability to hear some of the higher frequency sounds. These sounds are highly directional and help your brain determine which direction a sound is coming from. The loss of these higher frequencies also make it more difficult to distinguish sounds clearly in a crowded room. In addition, people with high-frequency hearing loss often complain they can hear the sounds but not understand the words.

10.3 Review

SUMMARY

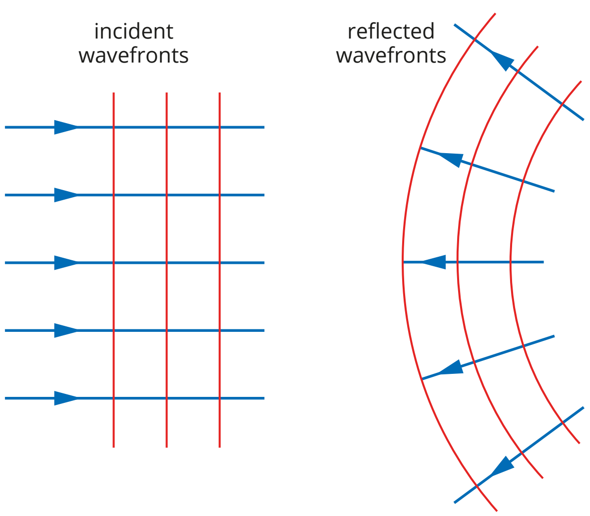
- A wave reaching the boundary between two materials in which it can travel will always be partly reflected, partly transmitted and partly absorbed.
- A wave has been reflected if it bounces back after reaching a boundary or surface.
- Waves reflect with a 180°, or $\frac{\lambda}{2}$, phase change from fixed boundaries. That is, crests reflect as troughs and troughs reflect as crests.
- Waves reflect with no phase change from free boundaries. That is, crests reflect as crests and troughs as troughs.
- When a wave is reflected from a surface, the angle of reflection will equal the angle of incidence.
- Refraction occurs when a wave changes speed as it passes from one medium to another.
- The refraction of the wave is measured by comparing the angle of incidence, i, to the angle of refraction, r.
- When a wave front speeds up at the boundary between two media, the wavelength increases and the wave front will bend away from the normal.
- When a wave front slows down at the boundary between two media, the wavelength decreases and the wave front will bend towards the normal.
- The frequency of the wave is unaffected.
- Total internal reflection occurs when the reflected angle is 90° to the normal.
- The incident angle at which total internal reflection occurs is called the critical angle.
- Diffraction, or spreading out of waves, around an obstacle or through an aperture occurs when λ is similar to or greater than the width of the aperture.
- Diffraction effects in sound allow sound to be heard around corners and are more prominent at low frequency.

KEY QUESTIONS

1 A wave travels along a rope and reaches the fixed end of the rope. What occurs next?

2 Which of the following properties of a wave can change when the wave is reflected: frequency, amplitude, wavelength or speed?

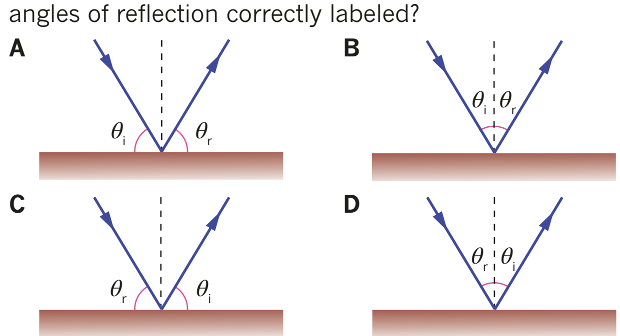
3 The following diagram shows a wave before and after being reflected from an object.



What is the shape of the object?

A flat **B** concave
C convex **D** parabolic

4 In which diagram are the angles of incidence and angles of reflection correctly labeled?



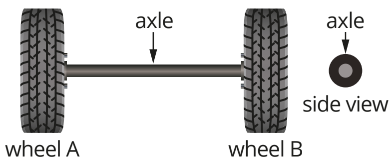
5 A geophysicist studying seismic waves determines that part of the interior of the Earth is molten. Explain why this statement can be made.

6 When water waves travelling through deep water reach a region of shallow water they can be refracted. Which of the properties of water waves given below always changes when the waves reach shallower water? (More than one correct answer is possible.)

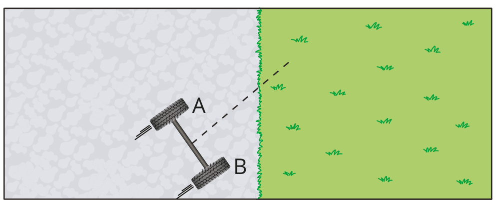
A frequency
B wavelength
C direction
D speed

10.3 Review *continued*

7 A popular analogy used to model refraction involves two wheels (A and B) attached to either end of an axle as shown in the figure below.



Rolling these wheels from a surface where they move faster (like on concrete) to where they move slower (like on grass) models the process of refraction (shown in the figure below).



Choose the correct words from those in brackets to complete the description of how this model works.

As wheel B rolls onto the grass it [speeds up/slows down]. Since wheel A is now moving [faster/slower] than wheel B, the wheels change [direction/speed]. When wheel A rolls onto the grass the wheels' direction [continues/stops] changing.

8 The velocity of sound increases by about 0.60 m s^{-1} for each degree increase in temperature. The speed of sound in air at 0.0°C is 331 m s^{-1}. A flute plays a note of 880 Hz at 20°C. A short distance above the ground there is a sudden increase in temperature to 30°C.

- **a** Calculate the speed of sound in air at 20°C.
- **b** Calculate the speed of sound in air at 30°C.
- **c** If the sound travels through cold air and is incident on the warm air at an angle of 50°, what will happen to the refracted angle?
- **d** Calculate the refracted angle.
- **e** At what angles must the sound be incident on the warm air layer for the sound to be reflected rather than refracted?

9 A flute and a tuba are being played at the same time. The flute is producing a note with a frequency of 2000 Hz and the tuba is producing a note with a frequency of 125 Hz at the same volume. A listener to the side of the auditorium complains that the tuba is drowning out the flute. How can this be so?

10.4 Wave interactions—superposition, interference and resonance

SUPERPOSITION AND INTERFERENCE

The sounds produced by musical instruments and the human voice are the product of the interaction of many waves. Imagine two transverse mechanical waves travelling towards each other along a string, as shown in Figure 10.4.1a. When the crest of one wave coincides with the crest of the other, the resulting displacement of the string is the vector sum of the two individual displacements (Figure 10.4.1b). The amplitude at this point is increased and the shape of the string resembles a combination of the two pulses. After they interact, the two pulses continue unaltered (Figure 10.4.1c). The resulting pattern is a consequence of the principle of **superposition**. The combination, or superposition, of the separate waves is called **interference**. In this case, as the displacement of the two waves was in the same direction, the two waves were added together, and **constructive interference** occurred.

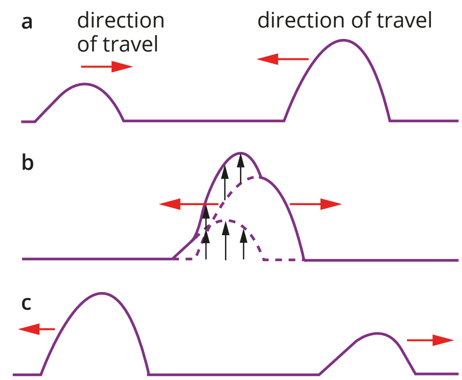


FIGURE 10.4.1 (a) As two wave pulses approach each other superposition occurs. (b) The occurrence of constructive interference. (c) After the interaction, the pulses continue unaltered; they do not permanently affect each other.

When a pulse with a positive displacement meets one with a negative displacement as shown in Figure 10.4.2b the resulting displacement of the string is the vector sum of the two individual displacements; in this case a negative displacement adds to a positive displacement to produce a wave of smaller magnitude. This is called **destructive interference**. Once again, the pulses emerge from the interaction unaltered (Figure 10.4.2c).

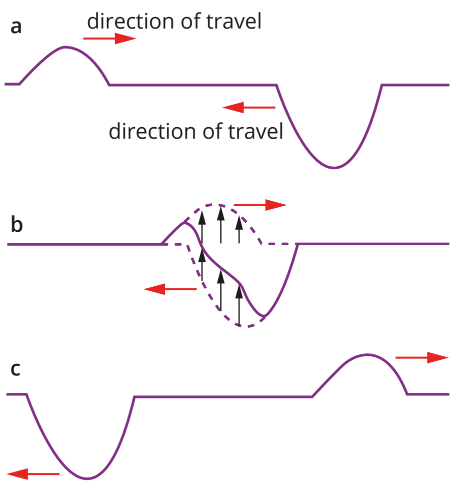


FIGURE 10.4.2 (a) As two wave pulses approach each other superposition occurs. (b) Superposition of waves in a string showing destructive interference. (c) As in constructive interference, the waves do not permanently affect each other.

When two waves meet and combine, there will be places where constructive interference occurs and places where destructive interference occurs. Although the wave pulses interact when they meet, passing through each other does not permanently alter the shape, amplitude or speed of either pulse. Just like transverse waves, longitudinal waves will also be superimposed as they interact. The terms superposition and interference are often used interchangeably.

A special case of **constructive interference** is where two waves of the same amplitude and wavelength, exactly **in phase**, add to together to give a wave of twice the amplitude as shown in Figure 10.4.3a. Complete **destructive interference** occurs when two waves are exactly opposite in phase (that is, one has a positive amplitude and the other a negative amplitude) and the two waves add together to give zero displacement (Figure 10.4.3b). The interference process can be used to explain the complex pattern seen in Figure 10.3.15b on page 356 where the interference or superposition of two diffraction patterns gives regions of complete destructive and constructive interference.

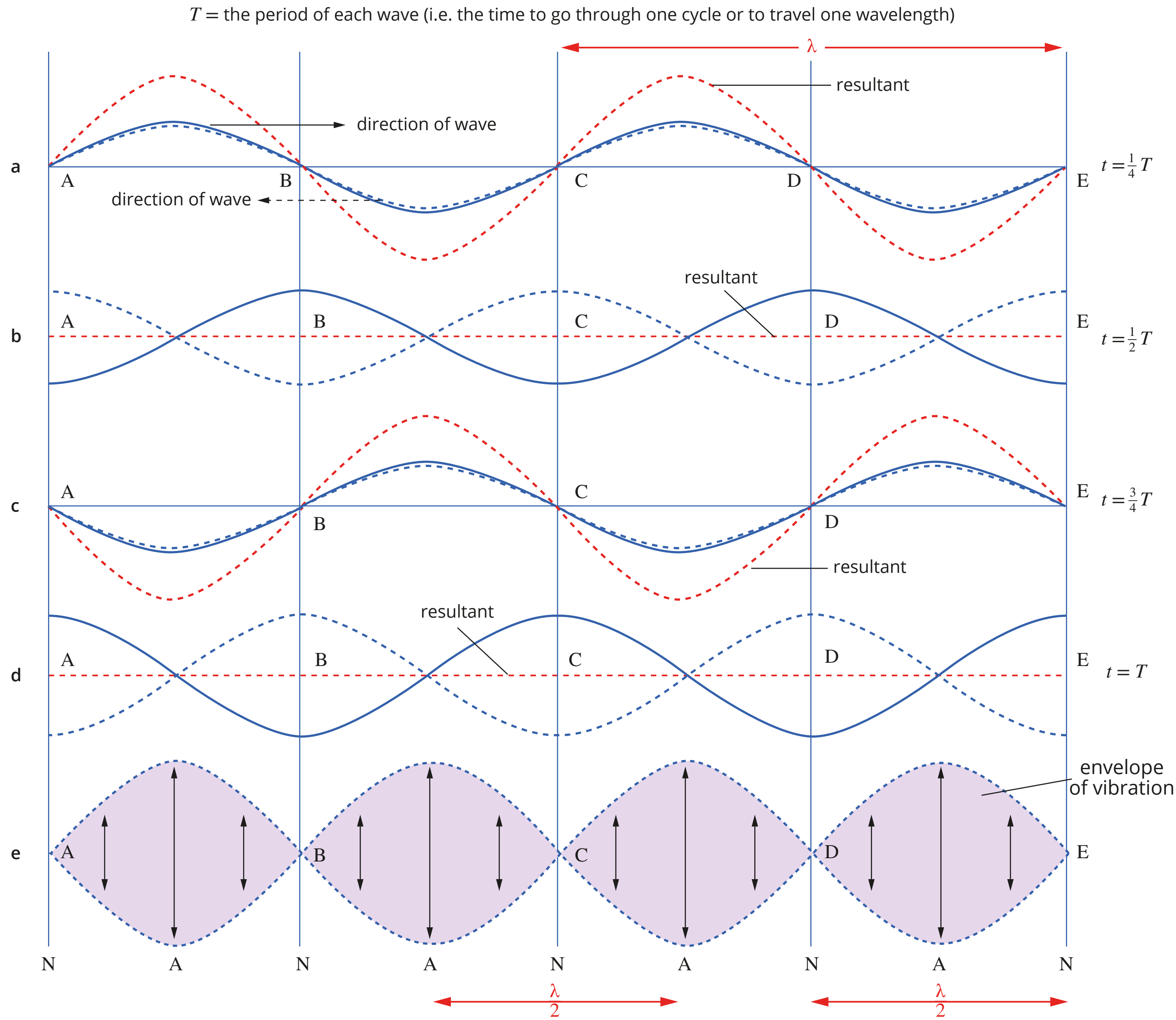


FIGURE 10.4.3 (a) Complete constructive interference occurs when two waves of the same amplitude, that are exactly in phase, add together to give a wave of twice the amplitude. (b) Complete destructive interference occurs when the two waves are 180° or $\frac{\lambda}{2}$ out of phase, add together to completely cancel out and give a wave of zero amplitude.

The effects of superposition and interference can be seen in many everyday examples. The ripples in the pond in Figure 10.4.4 were caused by raindrops hitting a pond. Where two ripples meet, a complex interference pattern is seen, where regions of constructive and destructive interference result from the superposition of the two waves. After this the ripples continue unaltered. Similarly, in a crowded room, all the sounds reaching your ear are superimposed, so that one complex sound wave arrives at the eardrum.

Superposition is important both theoretically and practically in the formation of complex sounds. Consider two single-frequency sound waves, or pure tones, one of which is twice the frequency of the other. The two individual waves are added together to give a more complicated resultant sound wave, as shown in Figure 10.4.5. Where one sound wave has a greater amplitude, as in the example illustrated, it will be the predominant sound heard. The quieter, higher frequency sound will combine with the louder one to create the sound that we hear. Note that for Figure 10.4.5, a transverse wave is used to depict the sound wave. The crests represent compressions (areas of high pressure) and the troughs represent rarefactions (areas of low pressure). The complex sounds that you emit when you speak or sing, or the square waves on a signal generator, can be modeled by superimposing various sine waves of different frequencies and amplitudes.

FIGURE 10.4.4 The ripples from raindrops striking the surface of a pond behave independently regardless of whether they cross each other or not. Where the ripples meet, a complex wave will be seen as the result of the superposition of the component waves. After interacting, the component waves continue unaltered.

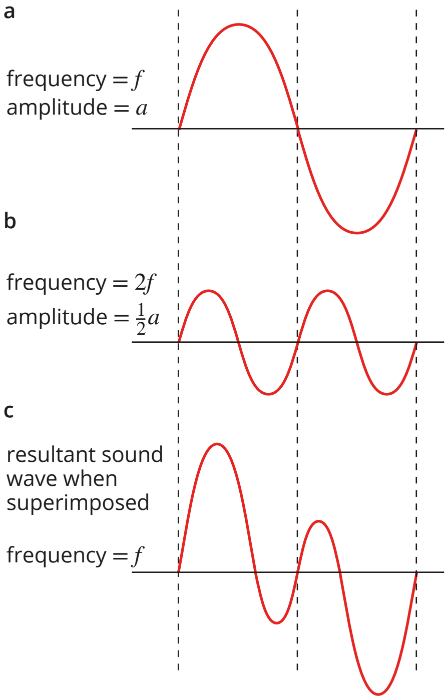


FIGURE 10.4.5 Two sound waves, one twice the frequency of the other, produce a complex wave of varying amplitude when they are superimposed.

PHYSICSFILE

The cocktail party effect and hearing loss

In a crowded room, individual sound waves will interfere with each other repeatedly, but it is still possible to distinguish which person is speaking. If you know the person's voice, then you know that their voice will sound the same. To discern one person's speech amid all the sounds in the room, your brain uses an innate ability to 'undo' the superposition of waves by selecting one person's voice and suppressing all the other noise. The 'cocktail party effect' also highlights the ability to hear your own name over the noise of a group of people talking.

PHYSICS IN ACTION

Effects of superposition and interference

Beats

If two sound waves of equal amplitude and slightly different frequency are added together, they superpose to form a resulting sound with a regular pulsation, known as a beat, as shown in Figure 10.4.6. At time X in the figure, the waves are in phase producing a maximum in the sound; then at time Y, the waves are out of phase, producing a minimum in the sound. The beat frequency is given by $f_{beat} = |f_2 - f_1|$. This phenomenon can be used to tune a musical instrument against a standard tone. It is also the reason why twin-engine aircraft often have an oscillating drone, and is due to the slightly different speeds of the two engines.

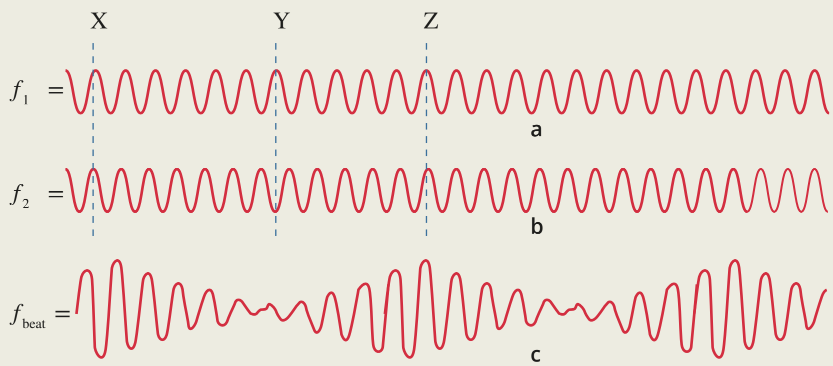


FIGURE 10.4.6 Superposition of two waves of slightly different frequencies f_1 and f_2 give rise to a wave with an oscillating amplitude.

Noise-cancelling headphones

Noise-cancelling headphones record the outside unwanted noise, and then produce a wave that is exactly opposite in phase, i.e. where there is a maximum in the amplitude of the recorded noise, there will be a minimum in the amplitude of the sound produced by the headphones. The superposition of the unwanted noise with the 'anti-noise' wave occurs over a time faster than your brain can process and thus creates destructive interference and effectively allows you to hear only the sound from your phone or device.

RESONANCE

You may have heard about singers who supposedly can break glass by singing particularly high notes. Figure 10.4.7 shows a glass being broken in this way. All objects that can vibrate tend to do so at a specific frequency known as their **natural frequency** or **resonant frequency**. **Resonance** is when an object is exposed to vibrations at a frequency equal to their resonant frequency. The vibration from one object causes a strong vibration in another. If the amplitude of the vibrations becomes too great, the object can be destroyed.

FIGURE 10.4.7 A glass can be destroyed by the vibrations caused by a singer emitting a sound of the same frequency as the resonant frequency of the glass.

A swing pushed once and left to swing or oscillate freely is an example of an object vibrating at its natural frequency. The frequency at which it moves backwards and forwards depends entirely on the design of the swing, mostly on the length of its supporting ropes. In time, the oscillations will fade away as the energy is transferred to the supporting frame and the air.

If you watch a swing in motion, it is possible to determine its natural oscillating frequency. It is then possible to push the swing at exactly the right time so that you match its natural oscillation. The additional energy you add by pushing will increase the amplitude of the swing rather than work against it. Over time, the amplitude will increase and the swing will go higher and higher; this is resonance. The swing can only be pushed at one particular rate to get this increase in amplitude (i.e. to get the swing to resonate). The frequency with which you push is the **forcing frequency**. If the rate you push is faster or slower, the forcing frequency that you are providing will not match the natural frequency of the swing and you will be fighting against the swing rather than assisting it.

Other examples of resonant frequency that you may have encountered are blowing air across the mouthpiece of a flute or drawing a bow across a string of a violin in just the right place (Figure 10.4.8). In each case, a clearly amplified sound is heard when the frequency of the forcing vibration matches a natural resonant frequency of the instrument.

FIGURE 10.4.8 The sound box of a stringed instrument is tuned to resonate for the range of frequencies of the vibrations being produced by the strings. When a string is plucked or bowed, the airspace inside the box vibrates in resonance with the natural frequency and the sound is amplified.

For resonance to occur, the forcing frequency must match the natural frequency. Two very significant effects occur when this happens.

- The amplitude of the oscillations within the resonating object will increase dramatically.
- The maximum possible energy from the source creating the forced vibration is transferred to the resonating object.

In musical instruments and loudspeakers, resonance is a desired effect. The sounding boards of pianos and the enclosures of loudspeakers are designed to enhance and amplify particular frequencies.

Resonance, however, can be undesirable in mechanical systems, such as car exhaust systems, it can also cause screws and ropes to loosen with spectacular and tragic consequences. On 16 April 1850, almost 200 soldiers died when a bridge in Angers, France, collapsed as a result of the resonance due to the oscillations caused by the column soldiers marching in step. Nowadays, care is taken to design mechanical systems that prevent resonance.

PHYSICSFILE

Tacoma Narrows Gorge suspension bridge

Resonance was responsible for destroying a suspension bridge over the Tacoma Narrows Gorge in the US State of Washington in 1940. Wind gusts of 70 km h^{-1} caused vibrations with a forcing frequency that caused the bridge to oscillate with ever-increasing amplitude, until the whole bridge shook itself apart. That is, the gusts of wind provided a forcing frequency that perfectly matched the natural oscillating frequency of the bridge. This caused the bridge to vibrate more and more until eventually it was destroyed. Figure 10.4.9 shows the bridge during oscillation and after collapse. You can find video clips of the bridge falling if you search for it online.

FIGURE 10.4.9 Original photos of the Tacoma bridge collapse. (a) The bridge oscillating due to resonance between the wind gusts and the natural frequency of the bridge. (b) The bridge after collapse.

10.4 Review

SUMMARY

- The principle of superposition states that when two or more waves interact, the resultant displacement or pressure at each point along the wave will be the vector sum of the displacements or pressures of the component waves.
- The process of superposition is also known as interference.
- Constructive interference occurs when the superposition of waves gives an increase in amplitude in the positive and/or negative direction.
- Complete constructive interference occurs when two waves of the same amplitude that are in phase combine to give a wave of double the amplitude.
- Destructive interference occurs when the superposition of two waves gives a reduction in amplitude.
- Complete destructive interference occurs when two waves of equal amplitude, exactly 180° or $\frac{\lambda}{2}$ out of phase, combine to give a wave with zero displacement.
- Resonance occurs when the frequency of a forcing vibration equals the natural or resonant frequency of an object.
- Two special effects occur with resonance:
  - the amplitude of vibration increases
  - the maximum possible energy from the source is transferred to the resonating object.

KEY QUESTIONS

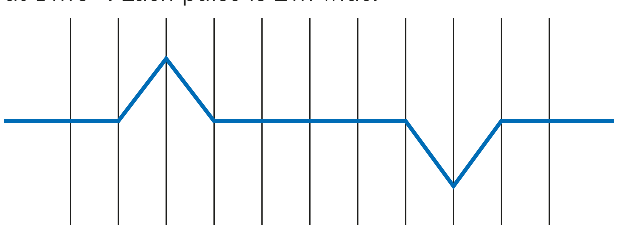
1 Which of the following about wave pulses are true and which are false? For the false statements, rewrite them so they are true.

a The displacement of the resultant pulse is equal to the sum of the displacements of the individual pulses.

b As the pulses pass through each other, the interaction permanently alters the characteristics of each pulse.

c After the pulses have passed through each other, they will have the same characteristics as before the interaction.

2 Two triangular wave pulses head towards each other at $1\,\text{m}\,\text{s}^{-1}$. Each pulse is 2 m wide.

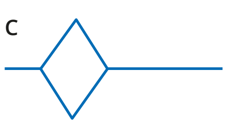


What will the superposition of these two pulses look like in 3 s?

A

B

C

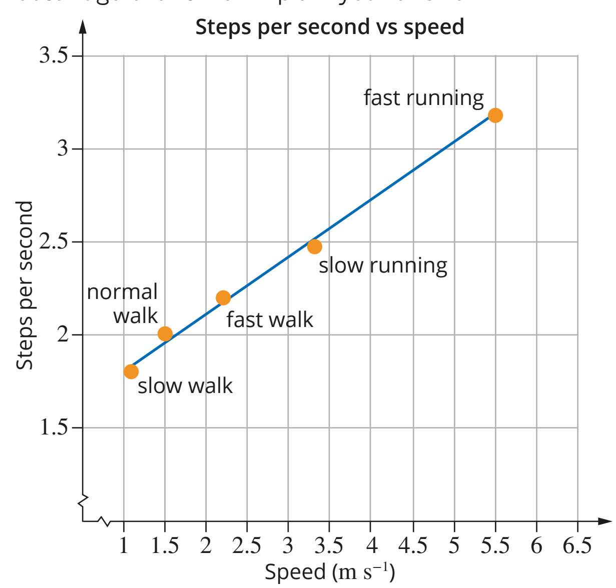


D

3 Explain why resonance can result in damage to man-made structures.

4 A light ray strikes a flat surface at an angle of 38° measured from the surface. What is the angle of reflection of the ray?

5 A footbridge over a river has a natural frequency of oscillation from side to side of approximately 1 Hz. When pedestrians walk at a pace that will produce an oscillation in the bridge close to the natural frequency of the footbridge, resonance will occur. The graph below displays relevant data about pedestrians walking or running. A pedestrian completes 1 cycle of their motion every 2 steps. Which activity of the pedestrians is most likely to cause damage to the footbridge over time? Explain your answer.



10.5 Standing waves and harmonics

Drawing a bow across a violin string causes the string to vibrate between the fixed bridge of the violin and the finger of the violinist (Figure 10.5.1). The simplest vibration will have maximum amplitude at the centre of the string, halfway between bridge and finger. This is a very simple example of a transverse **standing wave**.

FIGURE 10.5.1 Transverse standing waves can form along a violin string when the string is bowed by the violinist.

Standing waves are an important phenomenon of the superposition of waves. They occur when two waves of the same amplitude and frequency are travelling in opposite directions towards each other in the same string. Usually, one wave is the reflection of the other. Standing waves are responsible for the wide variety of sounds associated with speech and music.

STANDING WAVES IN A STRING

In Section 10.4 it was shown that when a wave pulse reaches a fixed end, it is reflected back 180° out of phase. That is, crests are reflected as troughs and troughs are reflected as crests.

Imagine creating a series of waves in a rope by shaking it vigorously. As the rope continues to be shaken, waves will travel in both directions. The new waves travelling down the rope will interfere with those being reflected back along the rope. This kind of motion will usually create quite a random pattern with the waves quickly dying away. Shaking the rope at just the right frequency, however, will create a new wave that interferes with the reflection in such a way that the two superimposed waves create a single, larger amplitude **standing wave**.

It is called a standing wave because the wave doesn't appear to be travelling along the rope. The rope simply seems to oscillate up and down with a fixed pattern. This situation contrasts with a standard travelling wave, in which a transverse wave is created where every point on the rope would have a maximum displacement at some time as the wave travels along the rope.

In Figure 10.5.2a–d, two waves (drawn in blue) are shown travelling in opposite directions towards each other along a rope. One of the waves is a string of pulses (shown as a solid line) and the other is its reflection (shown as a dashed line). The two waves superimpose when they meet. Since the amplitude and frequency of each is the same, the end result, as shown in part e, is a standing wave. At the points where complete destructive interference occurs, the two waves totally cancel each other out and the rope will remain still. These are called **nodes**. Where the rope oscillates with maximum amplitude, complete constructive interference is occurring. These points on the standing wave are called **antinodes**.

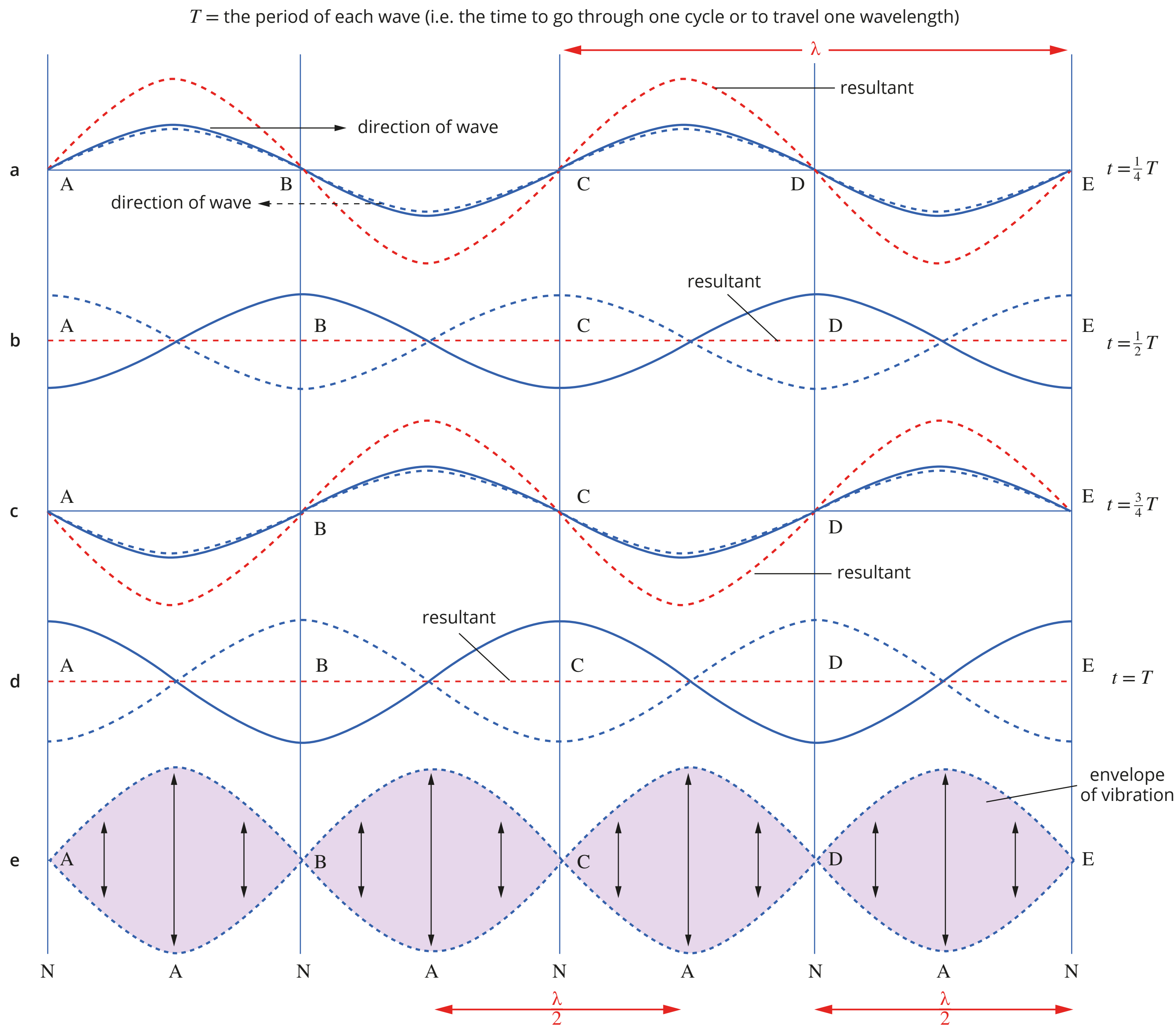


FIGURE 10.5.2 A standing wave created in a rope from two waves travelling in opposite directions, each with the same amplitude and frequency.

In Figure 10.5.2a $t = \frac{1}{4}T$: at a particular point in time, the two waves are completely superimposed (red dotted line), and complete constructive interference is occurring, resulting in a wave twice the original amplitude.

In Figure 10.5.2b $t = \frac{1}{2}T$: after a further period of time equal to $\frac{T}{4}$, the waves will each have moved $\frac{\lambda}{4}$, which means that they have moved a total of $\frac{\lambda}{2}$ in relation to each other. The waves are completely out of phase, complete destructive interference is occurring and the resulting displacement is zero.

As more time goes by, the waves will continue to move past each other.

In Figure 10.5.2c $t = \frac{3}{4}T$: the waves constructively interfere in the opposite direction and in Figure 10.5.2d $t = T$ the waves completely cancel again.

In Figure 10.5.2e the cycles shown in (a)–(d) form a standing wave. A standing wave swings between maximum displacements, creating antinodes (A) which lie halfway between the stationary nodes (N). Regardless of the position of the component waves, these nodes stay in the same place as the displacement at these points is always zero. Successive nodal points lie $\frac{\lambda}{2}$ apart, as do successive antinodal points.

Nodes and antinodes in a standing wave remain in a fixed position for a particular frequency of vibration. Figure 10.5.3 illustrates a series of possible standing waves in a rope, with both ends fixed, corresponding to three different frequencies. The lowest frequency of vibration, see in part a, produces a standing wave with one antinode in the centre of the rope. The ends are fixed so they will always be nodal points. Assuming the tension in the rope is kept the same, patterns b and c are produced at twice and three times the original frequency respectively.

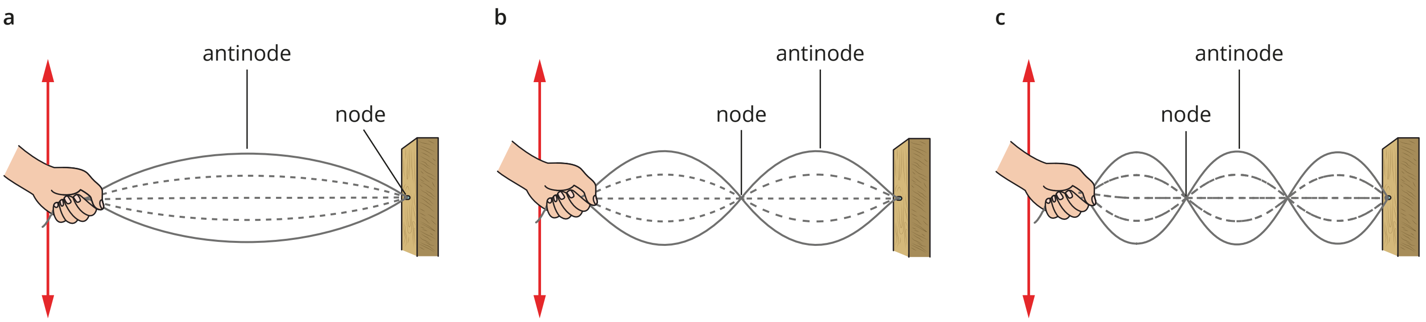


FIGURE 10.5.3 A rope vibrated at three different resonant frequencies, illustrating the standing waves produced at each frequency.

The rope could also vibrate at a frequency four times that of the original, and so on. The frequencies at which standing waves are produced are called the resonant frequencies of the rope.

It is important to note that the formation of a standing wave does not mean that the string or rope itself is stationary. It will continue to oscillate as further wave pulses travel up and down the rope. It is the relative position of the nodes and antinodes that remain unchanged.

It is also important to note that standing waves are not a natural consequence of every wave reflection.

Standing waves don't only exist in the everyday world but also occur at the subatomic level, and are the reason electrons exist in electron shells at fixed energies. This concept is covered in detail in Year 12.

Standing waves are only produced by the superposition of two waves of equal amplitude and frequency, travelling in opposite directions.

Standing waves are a result of resonance and occur only at the natural frequencies of vibration, or resonant frequencies, of the particular medium.

HARMONICS IN A STRING

Stringed musical instruments are examples of strings fixed at both ends. A large variety of waves of different frequencies/wavelengths are created. They travel along the string in both directions and reflect from the fixed ends, undergoing a phase change. Most of these vibrations will interfere in a random fashion and die away. However, those corresponding to resonant frequencies of the string will form standing waves and remain.

The resonant frequencies produced in this complex vibration of multiple standing waves are called **harmonics**. The lowest and simplest form of vibration, with one antinode (Figure 10.5.3a), is called the **fundamental** frequency. Higher-level harmonics (Figure 10.5.3b and c) are referred to by musicians as **overtones**.

The fundamental frequency usually has the greatest amplitude, so it has the greatest influence on the sound. The amplitude generally decreases for each subsequent harmonic. Usually all possible harmonics are produced in a string simultaneously, and the instrument and the air around it also vibrate to create the complex mixture of frequencies heard as an instrumental note.

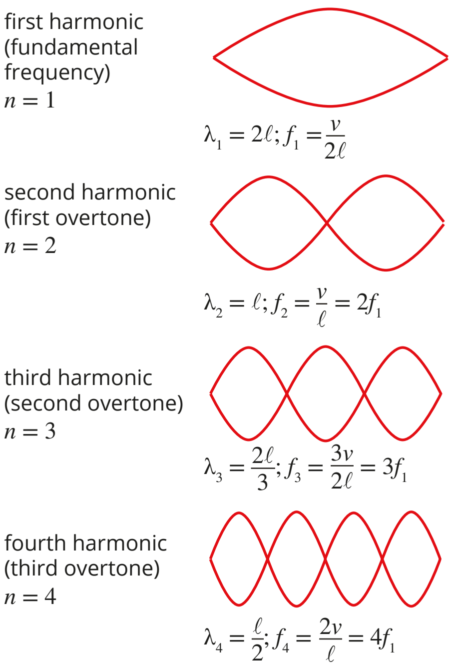


FIGURE 10.5.4 The first four resonant frequencies, or harmonics, in a stretched string fixed at both ends. The ends are fixed so they will always be nodal points.

The resonant frequencies or harmonics in a string of length l can be calculated from the relationship between the length of the string and the wavelength, λ, of the corresponding standing wave. Refer to Figure 10.5.4 to assist with understanding the following derivations.

For a string fixed at both ends:

The **first harmonic**, or fundamental frequency, has one antinode in the centre of the string and a node at each end, therefore there is only half a wavelength on the string

$$\ell = \frac{\lambda_1}{2} \text{ therefore } \lambda_1 = 2\ell$$

The second harmonic (first overtone) will have two antinodes and three nodes, therefore there is only one wavelength on the string.

$$\ell = \lambda_2 \text{ therefore } \lambda_2 = \ell = \frac{2\ell}{2}$$

The third harmonic (second overtone) will have three antinodes and four nodes, therefore there are $1\frac{1}{2}$ wavelengths on the string.

$$\ell = 3\frac{\lambda_3}{2} \text{ so } \lambda_3 = \frac{2\ell}{3}$$

And so in general, for any harmonic:

$$l = n\frac{\lambda_n}{2}$$

$$\lambda_n = \frac{2\ell}{n}$$

where λ_n is the wavelength (m)

ℓ is the length of the string (m)

n is the number of the harmonic, which is also the number of antinodes (i.e. 1, 2, 3, 4…)

Using the wave equation $v = f\lambda$ gives the relationship between frequency, velocity and string length.

For the first harmonic, or fundamental frequency:

$$\lambda_1 = 2\ell \text{ and } v = f_1\lambda_1 \text{ and so } f_1 = \frac{v}{\lambda_1} = \frac{v}{2\ell}$$

For the second harmonic (first overtone):

$$\lambda_2 = \ell \text{ and } v = f_2\lambda_2 \text{ so } f_2 = \frac{v}{\lambda_2} = \frac{v}{\ell} \text{ and } f_2 = 2f_1$$

For the third harmonic (second overtone):

$$\lambda_3 = \frac{2\ell}{3} \text{ and } v = f_3\lambda_3 \text{ so } f_3 = \frac{v}{\lambda_3} = \frac{3v}{2\ell} \text{ and } f_3 = 3f_1$$

And so in general:

$$f_n = \frac{nv}{2\ell} \text{ and } f_n = nf_1$$

where n is the number of the harmonic

f is the frequency of the wave (Hz)

v is the velocity of the wave (m s^{-1})

ℓ is the length of the string (m)

It should also be noted that the resonant frequencies of a string correspond to its tension and mass per unit length. Tightening or loosening the string will change the wavelengths and resonant frequencies for that string (i.e. the instrument will need tuning by adjusting the tension of the string). Heavier strings of a particular length will have different resonant frequencies than lighter strings of the same length and tension. For example, in a guitar the deeper notes are produced by the thicker strings.

Worked example 10.5.1

FUNDAMENTAL FREQUENCY

A violin string, fixed at both ends, has a length of 22 cm. It is vibrating at its fundamental frequency of vibration, at a frequency of 880 Hz.

a Calculate the wavelength of the fundamental frequency.	
Thinking	**Working**
Identify the length of the string (ℓ) in metres and the harmonic number (n).	$\ell = 22\,\text{cm} = 0.22\,\text{m}$ $n = 1$
Recall that for any frequency, $\lambda = \frac{2\ell}{n}$. Substitute the values from the question and solve for λ.	$\lambda = \frac{2\ell}{n}$ $= \frac{2 \times 0.22}{1}$ $= 0.44\,\text{m}$

b Calculate the wavelength of the second harmonic.	
Thinking	**Working**
Identify the length of the string (ℓ) in metres and the harmonic number (n).	$\ell = 22\,\text{cm} = 0.22\,\text{m}$ $n = 2$
Recall that for any frequency, $\lambda = \frac{2\ell}{n}$. Substitute the values from the question and solve for λ.	$\lambda = \frac{2\ell}{n}$ $= \frac{2 \times 0.22}{2}$ $= 0.22\,\text{m}$

Worked example: Try yourself 10.5.1

FUNDAMENTAL FREQUENCY

A standing wave in a string is found to have a wavelength of 0.50 m for the fundamental frequency of vibration. Assume that the tension in the string is not changed and that the string is fixed at both ends.

a Calculate the length of the string.

b Calculate the wavelength of the third harmonic.

PHYSICSFILE

Surface waves

Seismic surface waves travel along the boundary between materials, such as the Earth's crust and upper mantle. One type of surface wave is called the Rayleigh wave, or ground roll. These are surface waves that travel as ripples with a motion like that of waves on the surface of water, although the restoring force is elastic rather than gravitational as it is for water waves. A phenomenon known as free oscillation of the Earth is the result of the superposition between two such surface waves travelling in opposite directions creating a surface standing wave.

The first observations of free oscillations of the Earth were made during the 1960 Chile earthquake. Since then thousands of harmonics have been identified.

WIND INSTRUMENTS AND AIR COLUMNS

Longitudinal standing waves are also possible in air columns. These create the sounds associated with wind instruments. Blowing over the hole of a flute (Figure 10.5.5) or the reed of a saxophone produces vibrations that correspond to a range of frequencies that create standing waves in the tube.

FIGURE 10.5.5 Blowing over the mouthpiece of a flute and controlling the length of the flute with the keys enables a particular note to be produced.

The compressions and rarefactions of the sound waves, confined within the tube, reflect from both open and closed ends. This creates the right conditions for resonance and the formation of standing waves. The length of the pipe will determine the frequency of the sounds that will resonate.

The open end of a pipe corresponds to the fixed end of a string in that the reflected wave is fully inverted (i.e. undergoes a 180° change of phase). This means a compression is reflected as a rarefaction and a rarefaction is reflected as a compression. At a closed end, there is no change of phase, so a compression is reflected as a compression and a rarefaction as a rarefaction.

OPEN-ENDED AIR COLUMNS

Air columns open at both ends are referred to as 'open-ended' air columns. The flute is a typical example of a pipe open at both ends in which an air column can be made to vibrate. Another example includes the muffler in a car exhaust system.

A compression or rarefaction in a longitudinal sound wave will reflect from an open end with a phase change of $\frac{\lambda}{2}$.

Thus, the harmonics in the pressure variation for the wave will be very similar in nature to those of a string fixed at both ends.

Figure 10.5.6a shows the variation in pressure for the various harmonics formed in the pipe. The antinode, which is the region of maximum pressure variation, is shown as maximum positive and negative amplitude. The nodes, which are the regions of minimum pressure variation, are shown as zero displacement. Figure 10.5.6b shows the corresponding **particle displacement** for each of the pressure-variation harmonics.

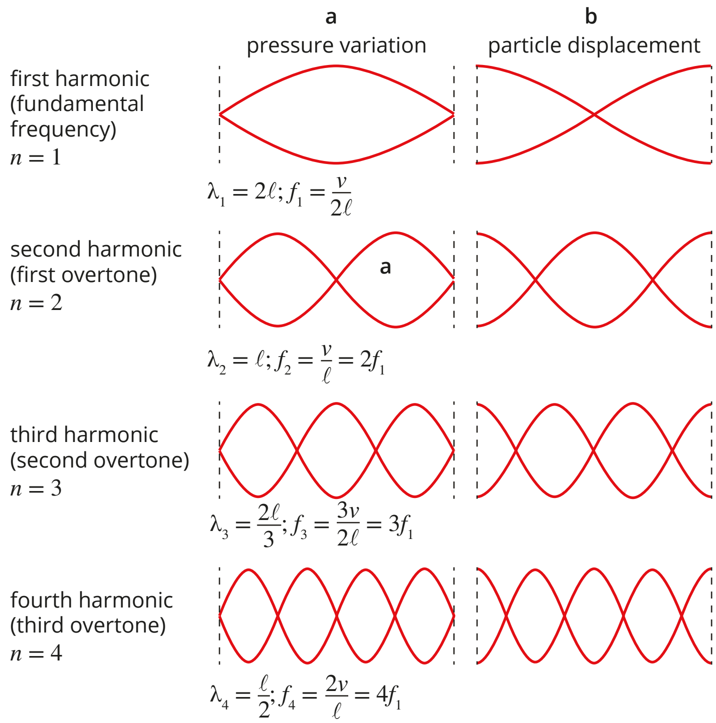


FIGURE 10.5.6 (a) the variation in pressure for each harmonic when standing waves are formed in a pipe open at both ends (b) The corresponding particle displacement for each harmonic.

Where there is an antinode in pressure variation, there is maximum pressure variation and therefore minimum particle displacement. Similarly where there is a node in pressure variation, there is an antinode (maximum) in particle displacement, as shown in Figure 10.5.7.

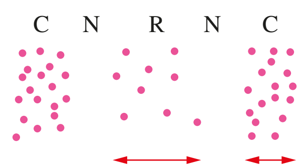


FIGURE 10.5.7 The diagram illustrates a sound wave showing compressions (C), rarefactions (R) and regions of normal pressure (N). The arrows indicate the particle displacement. In the regions of maximum pressure variation (C) and (R), the particle displacement will be at a minimum. In the region of minimum pressure variation (N), the particle displacement will be maximum.

An antinode in pressure variation corresponds to a minimum (node) in particle displacement.

A node in pressure variation corresponds to an antinode particle displacement.

When analysing the harmonics for the pressure wave in a pipe open at both ends, the same reasoning can be used as for a string fixed at both ends. Refer to Figure 10.5.6 throughout the following analysis.

$n = 1$: The **first harmonic**, or fundamental frequency, has one pressure antinode (compression) in the centre of the pipe and a node at each end, therefore there is only half a wavelength in the pipe.

$\ell = \frac{\lambda_1}{2}$ and $\lambda_1 = 2\ell$. Using $v = f_1\lambda_1$, then $f_1 = \frac{v}{\lambda_1} = \frac{v}{2\ell}$.

Conversely the particle displacement has a maximum (antinode) at the end of the pipe and a minimum (node) in the centre.

$n = 2$: The second harmonic or first overtone will have two pressure antinodes and three nodes, therefore there is only one wavelength in the pipe.

$\ell = \lambda_2$, therefore $\lambda_2 = \ell = \frac{2\ell}{2}$. Using $v = f_2\lambda_2$, then $f_2 = \frac{v}{\lambda_2} = \frac{v}{\ell}$ and $f_2 = 2f_1$.

The particle displacement has three antinodes—one at each end and one in the middle—and two nodes.

$n = 3$: The third harmonic or second overtone will have three pressure antinodes and four nodes, therefore there are $1\frac{1}{2}$ wavelengths in the pipe.

$\ell = 3\frac{\lambda_3}{2}$ so $\lambda_3 = \frac{2\ell}{3}$. Using $f_3 = \frac{v}{\lambda_3} = \frac{3v}{2\ell}$ and $f_3 = 3f_1$.

The particle displacement has three nodes and four antinodes.

And so in general,

For a harmonic in a pipe open at both ends

$$\ell = n\frac{\lambda_n}{2}$$

$$\lambda_n = \frac{2\ell}{n}$$

where λ is the wavelength (m)

ℓ is the length of the string (m)

n is the number of the harmonic, which is also the number of pressure antinodes (i.e. 1, 2, 3, 4…)

It should also be noted that, due to the air pressure from the sound wave in the tube, the standing wave produced will extend slightly beyond the end of the tube. The actual length that should be used will therefore be a little longer than the tube itself. The distance that the sound wave will extend beyond the end of the tube depends partially on the diameter of the tube itself and also on the surrounding air pressure. Since the discrepancy is small, for the purposes of this study the assumption will be made that the length of the tube coincides with the length of the standing wave.

Worked example 10.5.2

OPEN-ENDED AIR COLUMNS

A particular flute has an effective length of 35 cm. It can be thought of as an open-ended air column. The speed of sound is 340 m s^{-1}.

a Calculate the wavelength of the second harmonic.	
Thinking	**Working**
Identify length of the air column (ℓ) in metres and the harmonic number (n).	$\ell = 35$ cm $= 0.35$ m n = 2
Recall that for any frequency, $\lambda_n = \frac{2\ell}{n}$. Substitute the values from the question and solve for λ.	$\lambda_n = \frac{2\ell}{n}$ $\lambda_2 = \frac{2 \times 0.35}{2}$ $= 0.35$ m

b Determine the frequency of the second harmonic.	
The frequency can be calculated using $f_2 = \frac{v}{\lambda_2}$	$f_2 = \frac{340}{0.35}$ $= 971$ Hz

c Determine the frequency of the fourth harmonic.	
Recall $f_n = nf_1$	$f_4 = 4f_1$ $= 2f_2$ $= 2 \times 971$ $= 1940$ Hz

PHYSICSFILE

Wind Instruments

Wind instruments do not have the benefit of having strings of different masses that can be tensioned or released to produce different frequencies. In bugles or ceremonial trumpets, where the length is fixed, the pitch of the note produced must come from the range of harmonics available as a direct result of that fixed length. This severely restricts the number of notes that can be played.

Other instruments allow the player to vary the length of the pipe by covering or uncovering holes along the length of the pipe or, in the case of orchestral trumpets, using valves to connect additional curved lengths of pipe. The coiled lengths of pipe in trombones, French horns and trumpets allow a greater total length and thus a greater potential range of notes. Part of the skill of a woodwind or brass musician is controlling which mode of vibration dominates the final sound.

Worked example: Try yourself 10.5.2

OPEN-ENDED AIR COLUMNS

The wavelength of the fourth harmonic in a tube that can be considered as an open-ended air column is found to be 12 cm.

a Calculate the length of the tube, assuming that the standing wave does not extend beyond the ends of the tube.

b Determine the fundamental frequency.

Air columns closed at one end

In this context, a 'closed air column' means that the pipe or tube is closed at one end and remains open at the other. Some examples of closed air columns are the human vocal tract, the ear canal, ported loudspeakers and car engine manifolds.

In both strings and fully open pipes, the reflection of the wave is the same at both ends. The *open end* of the air column reflects a sound wave with a change of phase. However, in a pipe closed at one end it's different. At the *closed end*, there will be no change of phase for the reflected sound wave. Here the reflected waves will interfere constructively with the incoming waves so there will be a maximum in pressure variation (a compression). However, the air particle movement (displacement) will be minimal in this region.

The result is that the pressure variation established in a tube closed at one end will have a node at the open end and an antinode at the closed end, as shown in Figure 10.5.8a. Conversely, there will be a maximum in particle displacement at the open end and a minimum at the closed end, as shown in Figure 10.5.8b

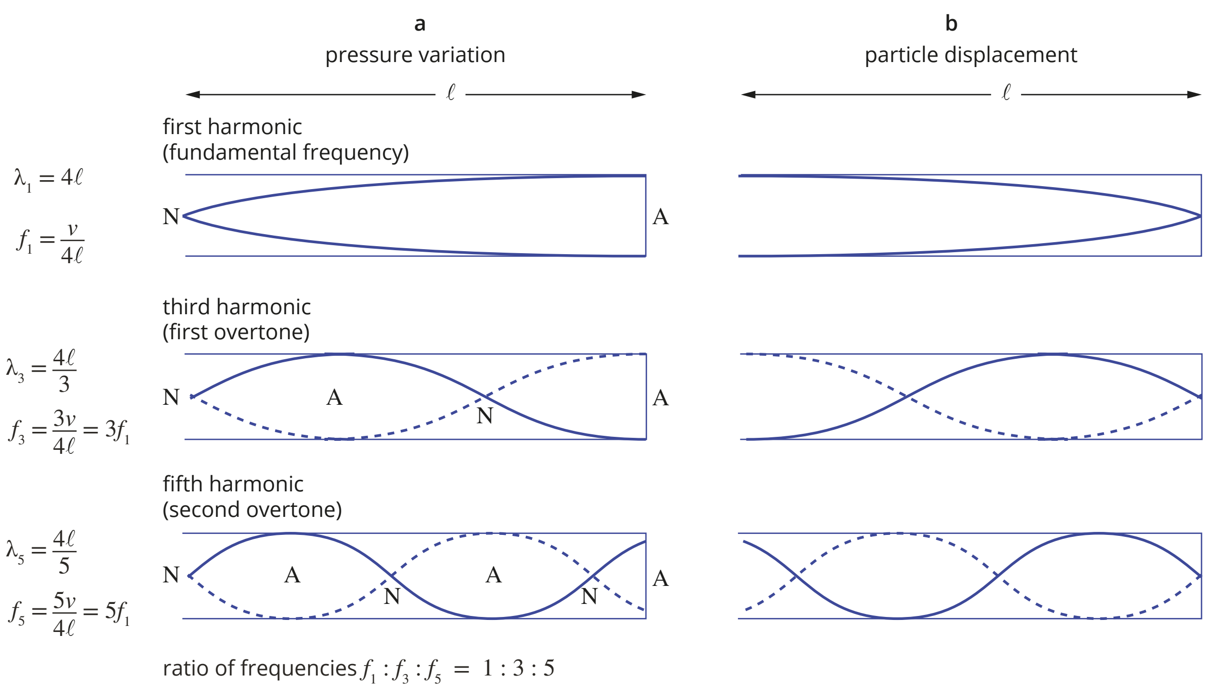


FIGURE 10.5.8 (a) The lower harmonics for a pipe closed at one end. Only odd-numbered harmonics are possible, since only these satisfy the condition of having a node in pressure variation at the open end and an antinode at the closed end. (b) The equivalent particle displacement. The particles have maximum displacement at the open end and minimum displacement at the closed end.

For a tube closed at one end the pressure variation will have an antinode at the closed end and a node at the open end.

Particle displacement will be a minimum at the closed end and a maximum at the open end.

For the following analysis refer to Figure 10.5.8a.

For the first harmonic, or fundamental frequency, only a quarter of a wavelength will fit into the air column, therefore it will have a wavelength four times the length of the effective air column.

The first harmonic ($n = 1$), or fundamental frequency, will have:

$\ell = \frac{\lambda_1}{4}$ and $\lambda = 4\ell$. Using $f = \frac{v}{\lambda}$ then $f_1 = \frac{v}{4\ell}$

$n = 2$ cannot form because this would result in an antinode at each end.

For the next harmonic three quarters of a wavelength can fit in the air column; therefore it will have a wavelength $\frac{4}{3}$ the length of the pipe. This is the $n = 3$ harmonic.

$\ell = \frac{3\lambda_3}{4}$ gives $\lambda_3 = \frac{4\ell}{3}$ and $f_3 = \frac{v}{\lambda_3} = \frac{3v}{4\ell}$

The next harmonic that can form will have five quarters of a wavelength that will fit in the air column, therefore:

$\ell = \frac{5\lambda_5}{4}$ and $\lambda_5 = \frac{4\ell}{5}$. This is the $n = 5$ harmonic.

In general, for a pipe closed at one end:

$$\ell = \frac{n\lambda_n}{4}$$

and

$$\lambda_n = \frac{4\ell}{n}$$

where: λ is the wavelength in metres (m)

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

n is the number of the harmonic; odd-number integers only (i.e. 1, 3, 5...)

The frequency of the harmonic is given by:

$$f_n = \frac{nv}{4\ell}$$

where n is the number of the harmonic (i.e. 1, 3, 5…)

Notice that only odd-numbered harmonics will form since the conditions necessary for a standing wave to form are only met when there is an antinode at the closed end of the pipe and a node at the open end. The ratio of the wavelengths of the harmonics will be 1 : 3 : 5 : ... That means the wavelength of the third harmonic is $\frac{1}{3}$ the length of the fundamental frequency (first harmonic), the fifth harmonic is $\frac{1}{5}$, and so on.

As has been noted the even numbered harmonics (i.e. 2, 4, 6...) will not form. The formula can be expressed alternatively in the following form.

$$\lambda_{2n-1} = \frac{4l}{2n-1} \text{ and } f_{2n-1} = \frac{(2n-1)v}{4l}$$

where λ_{2n-1} is the wavelength in metres (m)

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

n is an integer number (i.e. 1, 2, 3, 4 ...) and refers to the next harmonic in the sequence, not the harmonic number.

$2n-1$ refers to the harmonic number. For $n = 1$ this refers to the first harmonic, but for $n = 2$ this $((2 \times 2) - 1 = 3)$ is the 3rd harmonic.

As in open columns, the air pressure from the sound wave in the tube will cause the standing wave produced to extend slightly beyond the end of the tube. The actual length that should be used will therefore be a little longer than the tube itself. Since the discrepancy is small, for the purposes of this study the assumption will be made that the length of the tube coincides with the length of the standing wave.

Worked example 10.5.3

AIR COLUMN CLOSED AT ONE END

The ear canal, from the outer ear to the eardrum, can be thought of as a tube closed at one end (by the eardrum) and open at the other. It is approximately 3.0 cm long in an adult.	
a Calculate the wavelength of the fundamental frequency of the adult ear canal.	
Thinking	**Working**
Identify the length of the air column (ℓ) in metres and the harmonic number (n).	$\ell = 3$ cm $\ell = 0.03$ m $n = 1$
Recall that for any frequency for closed tubes, $\lambda_n = \frac{4\ell}{n}$. Substitute the values from the question and solve for λ.	$\lambda_n = \frac{4\ell}{n}$ $= \frac{4 \times 0.03}{1}$ $= 0.12$ m

b Assuming that the speed of sound through air is 340 m s^{-1}, what frequency does this wavelength correspond to?	
Thinking	**Working**
Identify the speed of the sound (v) in m s^{-1} and the wavelength (λ) from the previous question.	$v = 340$ m s^{-1} $\lambda_1 = 0.12$ m
Recall the wave equation $v = f\lambda$. Rearrange to find f.	$f = \frac{v}{\lambda_1}$ $= \frac{340}{0.12}$ $= 2833$ Hz ≈ 3 kHz

This answer explains why sounds of around 3 kHz are generally heard best by adult humans. The length of the ear canal means that sounds of around this frequency are best amplified by resonance forming a standing wave. There are, however, many other factors which influence the range of frequencies a person is able to hear.

Worked example: Try yourself 10.5.3

AIR COLUMN CLOSED AT ONE END

An air column closed at one end is 12 cm long. Assume that the standing wave does not extend beyond the end of the tube.

a Calculate the wavelength of the fifth harmonic.

b Calculate the wavelength of the third harmonic.

While the discussion in this section has been of two-dimensional standing waves, standing waves may also form in three dimensions, such as in a section of the Earth's crust. Standing waves will form as a result of resonance in any wave form.

PHYSICS IN ACTION

Music and musical scales

Musical instruments generally produce standing waves that are whole-number multiples of a fundamental frequency. A superposition of these waves generally produces a sound that is harmonious. Harmonies are an important part of music. For example, a ratio of 2:1 in frequency is called an octave. Middle C has a frequency of 261.6 Hz. If you go up an octave, the next C is at 523.3 Hz; i.e, double the frequency of middle C.

10.5 Review

SUMMARY

- Standing or stationary waves occur as a result of resonance at the natural frequency of vibration.
- Standing waves are produced by the superposition of waves of equal amplitude and frequency/wavelength travelling in the opposite direction.
- Points on a standing wave that remain still are called nodes.
- Points of maximum vibration (in a string) or pressure variation (in a tube) on a standing wave are called antinodes.
- The standing wave frequencies are referred to as harmonics. The simplest mode is referred to as the fundamental frequency.
- The harmonics above the fundamental frequency are also the called the first overtone, second overtone, third overtone and so on.
- Within a string fixed at both ends, the wavelength of the standing waves corresponding to the various harmonics is:

$$\lambda = \frac{2\ell}{n}$$

and the frequency is:

$$f = \frac{nv}{2\ell}$$

All harmonics may be present.

- Longitudinal standing waves can also form in air columns.
- At the open end a phase change occurs and a pressure variation node forms.
- If a pipe is closed at one end, there is no phase change and a pressure variation antinode is formed at the closed end.
- An antinode in the standing wave (maximum pressure variation) corresponds to a minimum in particle displacement
- A node in the standing wave (minimum pressure variation) corresponds to a maximum in particle displacement.
- For a pipe open at both ends, the pressure variation is similar to that of a string fixed at both ends, therefore:

$$\lambda_n = \frac{2\ell}{n}$$

and the frequency is:

$$f_n = \frac{nv}{2\ell}$$

- For a pipe with one end closed, an antinode forms at the closed end and a node forms at the open end:

$$\lambda_n = \frac{4\ell}{n}$$

and the frequency is:

$$f_n = \frac{nv}{4\ell}$$

- Expressed this way, only the odd numbered harmonics, $n = 1, 3, 5\ldots$, can be formed.

KEY QUESTIONS

1 A transverse standing wave is produced using a rope. Is the standing wave actually standing still? Explain your answer.

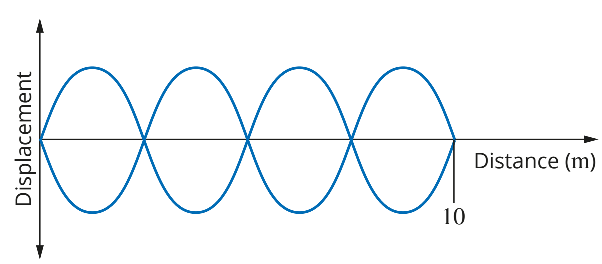
2 Describe how superposition and interference are related to the formation of standing waves in a stretched slinky spring.

3 What is the wavelength of the fundamental mode of a standing wave on a string 0.4 m long and fixed at both ends?

4 Calculate the length of a string fixed at both ends when the wavelength of the fourth harmonic is 0.75 m.

5 A standing wave is produced in a rope fixed at both ends by vibrating the rope with four times the frequency that produces the fundamental or first harmonic. How much larger or smaller is the wavelength of this standing wave compared to that of the fundamental or first harmonic?

6 A standing wave pattern in a string is shown over a distance of 10 m.



What is the length of the rope that would generate the first harmonic if a standing wave of the same wavelength shown in the diagram above was produced?

7 The fundamental frequency of a violin string is 350 Hz and the velocity of the waves along it is 387 m s^{-1}. What is the wavelength of the new fundamental when a finger is pressed to shorten the string to $\frac{2}{3}$ its original length?

8 A metal string (at constant tension) of length 50 cm is plucked, creating a wave pulse. The speed of the transverse wave created is 300 m s^{-1}. Both ends of the string are fixed.

   a Calculate the frequency of the fundamental frequency.

   b Calculate the frequency of the second harmonic.

   c Calculate the frequency of the third harmonic.

9 A flute can be considered an open-ended air column. For a flute of effective length 45.0 cm:

   a What is the wavelength of the fundamental frequency?

   b What is the wavelength of the second harmonic?

   c Calculate the frequency of the third harmonic, assuming the speed of sound in the flute is 330.0 m s^{-1}.

10 An organ pipe is an air column closed at one end with an effective length of 75.0 cm. The speed of sound inside the pipe is 330 m s^{-1}.

   a Calculate the frequency of the fundamental note produced by the pipe.

   b Calculate the frequency of the third harmonic.

   c Calculate the frequencies of the next two harmonics, after the third, that the pipe can produce.

11 **a** A pipe produces a fundamental frequency of 450 Hz and subsequent resonant frequencies of 900 Hz, 1350 Hz and 1800 Hz. Is this an open ended or closed pipe? Explain why.

   b A pipe produces a fundamental wavelength of 3.0 m and subsequent resonances at 1.0 m and 0.6 m. Is this an open-ended or closed pipe? Explain why.

10.6 Wave intensity and applications of wave properties

INTRODUCTION

A stone thrown into a pond produces a series of spreading ripples, and as these ripples move away from the point where the stone landed, their amplitude decreases. The energy of the original source is spread over a greater area. As a listener moves away from a sound source, a similar thing happens. The apparent loudness decreases and the sound becomes increasingly difficult to hear (Figure 10.6.1). Cupping your hands around your mouth, or using a simple megaphone, will limit the spread of the sound to some extent and allow it to be heard at a greater distance.

All of these things relate to the wave intensity experienced, which varies with the distance from the original source of the wave.

PHYSICSFILE

Tsunamis

A tsunami is usually generated by an underwater earthquake, and contains an enormous amount of energy. As it travels towards the coast, the water depth significantly decreases. However, a tsunami does not lose much of its energy as it travels due to its long wavelength, and therefore its height increases significantly as it approaches the coast.

FIGURE 10.6.1 As you get further from the source of a sound, the energy of the wave is distributed over a greater area so the intensity of the sound decreases. This means the sound isn't as loud.

CHANGE OF INTENSITY WITH DISTANCE

You will remember from earlier modules that waves carry energy. The **intensity** of a wave measures the amount of energy that passes through a square metre per second. The intensity of a wave is proportional to the square of the amplitude and is measured in $W\,m^{-2}$ (watts per square metre). You may recall from the studies of energy that 1 watt = 1 joule per second. A human ear can detect sounds ranging from $10^{-12}\,W\,m^{-2}$ to $1\,W\,m^{-2}$.

A wave will spread out spherically from a point source, as shown in Figure 10.6.2. When the distance from the source is doubled, the energy carried by the wave will be spread over four times the original area. This means that the intensity of the wave at any particular point will be reduced to a quarter of that at the original distance. At three times the original distance from the source, the intensity will be reduced to one-ninth of that at the starting point.

This is referred to as an **inverse square law**. If the source is considered to be a point source, then the wave intensity, I, will decrease with the square of the distance, r, from the source:

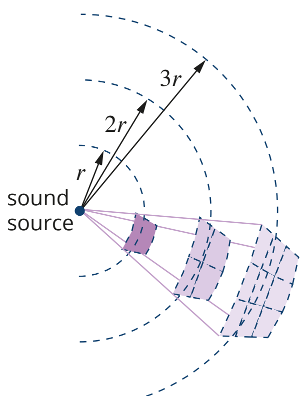


FIGURE 10.6.2 As it moves away from a source, a sound wave spreads radially. The same energy must spread over an area increasing with the square of the distance from the source, resulting in the intensity decreasing by the same ratio.

$$\ell \propto \frac{1}{r^2}$$

where ℓ is the intensity in watts per square metre ($W\,m^{-2}$)

r is the distance from the source in metres (m)

Based on this relationship, at double the distance, $I = \frac{1}{2^2} = \frac{1}{4}$ of the original intensity.

At four times the distance, $I = \frac{1}{4^2} = \frac{1}{16}$ of the original intensity, and so on.

The decrease in intensity follows this rule only if there is no other loss of energy within the medium. The decrease in intensity with distance can often be much higher when the material through which the wave is travelling absorbs the energy of the wave. Wind, temperature and other differences affecting the medium can also influence how the intensity changes with distance.

If the intensity is I_0 at a particular distance r_0, then $I_0 \propto \frac{1}{{r_0}^2}$.

At another distance r_f, the intensity is $I_f \propto \frac{1}{{r_f}^2}$.

Since the proportionality constants are the same, then the ratio of the intensities is given by $\frac{I_f}{I_0} = \frac{\frac{1}{{r_f}^2}}{\frac{1}{{r_0}^2}}$, which can be simplified to $\frac{I_f}{I_0} = \frac{{r_0}^2}{{r_f}^2}$.

Notice the values for distance are inverted compared with the intensity due to the inverse relationship.

Worked example 10.6.1

INTENSITY AND DISTANCE 1

Mia hears a siren sounding 10 m away. By what factor would the intensity of the sound from the siren change if Mia was to move to a distance of 50 m from the siren?

Thinking	Working
Intensity, I, will decrease with the square of the distance, r, from the source. The ratio of the intensity at 50 m to the original intensity at 10 m is the factor required. Identify the initial values I_0 and r_0 and the final values I_f and r_f.	$r_0 = 10\,\text{m}$ $I_0 \propto \frac{1}{{r_0}^2}$ then $I_0 \propto \frac{1}{10^2}$ $r_f = 50\text{ m}$ $I_f \propto \frac{1}{{r_f}^2}$ then $I_f \propto \frac{1}{50^2}$
Determine the relationship between the variables.	$\frac{I_f}{I_0} = \frac{\left(\frac{1}{{r_f}^2}\right)}{\left(\frac{1}{{r_0}^2}\right)}$ $\frac{I_f}{I_0} = \frac{{r_0}^2}{{r_f}^2}$
Evaluate.	$\frac{I_f}{I_0} = \frac{10^2}{50^2}$ $\frac{I_f}{I_0} = 0.04$

Worked example: Try yourself 10.6.1

INTENSITY AND DISTANCE 1

Sam heard an annoying sound from 100 m away. By what factor would the intensity of the annoying sound change if Sam was to move to a distance of 400 m from the sound?

PHYSICSFILE

Inverse square laws

The law for the intensity of light around a point source of light is an example of an inverse square law. There are a number of important inverse square laws in physics.

Some others you will encounter in physics are: Coulomb's law for the force between electric charges, and Newton's law of universal gravitation, for the gravitational force between any two masses.

In all these cases, something can be imagined to be spreading out evenly from a point. The intensity of this 'something' at a certain distance will therefore be inversely proportional to the area of the sphere over which it is spread. As the formula for the surface area of a sphere is $A = 4\pi r^2$, the intensity will therefore decrease with the square of the distance.

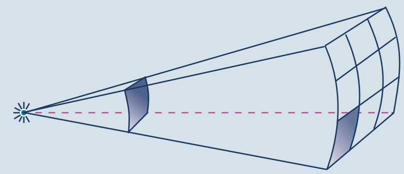


FIGURE 10.6.3 The inverse square law for a light source. At a distance of 1 m a square is drawn perpendicular to the rays. At a distance of 3 m from a point source, the light will be spread over an area nine times as large as that at 1 m. The light will therefore appear only one-ninth as bright.

Worked example 10.6.2

INTENSITY AND DISTANCE 2

An ambulance started its siren 200 m after leaving the scene of an accident. After a short time the intensity of the sound was measured as being a quarter of the original. Assuming the volume of the siren hadn't changed, how far away was the ambulance when the intensity was measured?

Thinking	Working
Intensity, I, will decrease with the square of the distance, r, from the source. The expression $\frac{I_f}{I_0} = \frac{r_0^2}{r_f^2}$ can be used. Identify the variables r_0 and $\frac{I_f}{I_0}$.	$r_0 = 200\,\text{m}$ $\frac{I_f}{I_0} = \frac{1}{4}$
Rearrange the expression and evaluate.	$\frac{I_f}{I_0} = \frac{r_0^2}{r_f^2}$ $r_f^2 = \frac{r_0^2 I_0}{I_f}$ $r_f^2 = \frac{200^2}{0.25} = 200^2 \times 4 = 160\,000$ $r_f = 400\,\text{m}$

Worked example: Try yourself 10.6.2

INTENSITY AND DISTANCE 2

A fog horn was originally heard from a boat when the boat was 1 km from the fog horn. After some time, the intensity of the fog horn was measured as being half of the original. Assuming the volume of the fog horn hadn't changed, how far away was the boat from the fog horn when the intensity was measured?

EXTENSION

The decibel scale

The human ear can detect sound intensities from $10^{-12}\,\text{W}\,\text{m}^{-2}$ to $1\,\text{W}\,\text{m}^{-2}$. Therefore, a sound intensity of $10^{-11}\,\text{W}\,\text{m}^{-2}$ is 10 times larger than an intensity of $10^{-12}\,\text{W}\,\text{m}^{-2}$. For everyday use sound levels are usually measured in decibels, which is a logarithmic scale where:

$$I(\text{dB}) = 10\log\left(\frac{I}{I_0}\right)$$

where I_0 is the threshold intensity of hearing ($10^{-12}\,\text{W}\,\text{m}^{-2}$)

I is the sound intensity.

Some typical values for sound intensity are given in Table 10.6.1. When viewing the table, bear in mind every 10 dB means 10 times the sound intensity.

The long-term effect on a person's hearing also depends on how long a person is exposed to loud noise. It is an occupational health and safety requirement that people who regularly work in a loud environment wear ear protection.

TABLE 10.6.1 Examples of common sounds, their sound intensity in $\text{W}\,\text{m}^{-2}$, the relative sound intensity compared to the threshold of hearing ($I_0 = 10^{-12}$), and the equivalent sound intensity in dB.

Sound	Sound intensity, I ($\text{W}\,\text{m}^{-2}$)	I/I_0	I (dB)
hearing threshold (I_0)	10^{-12}	$10^0 = 1$	0
whisper	10^{-10}	10^2	20
normal conversation	10^{-6}	10^6	60
heavy traffic	10^{-5}	10^7	80
loud truck close by	10^{-3}	10^9	90
fighter jet at 300 m	10^{-2}	10^{10}	100
front row of a rock concert, petrol-powered tools	10^{-1}	10^{11}	110
threshold of pain	10^1	10^{13}	130

APPLICATIONS OF WAVES—ULTRASOUND

Sound is an example of energy travelling in the form of a wave. Humans can hear sound waves in the frequency range 20 to 20 000 Hz. Frequencies above 20 000 Hz are called ultrasound.

Ultrasounds are used in medicine in two different ways: as a diagnostic tool and for treatment. The most common use of ultrasound is diagnostic, in the production of images of internal parts of the body. For example, ultrasound can be directed through the amniotic fluid within the womb to build a picture of a developing foetus. Diagnostic applications rely on the reflection of sound waves. During diagnostic applications of ultrasound, low-intensity waves are used. This limits any effects on the cells of the body.

Ultrasound can also be used as a therapy, ranging from the treatment of kidney stones to allowing surgeons to perform delicate brain surgery. The therapeutic application of ultrasound uses high-intensity sound waves on the cells of the body. Recall that sound waves are just mechanical vibrations in the particles of the carrying medium. During ultrasound procedures vibrations are created in the cells of the body. At low intensities, these vibrations are hardly more energetic than they would normally be. However, if the sound waves are intense enough, they can heat regions deep inside the body and damage or destroy cells.

Heat treatments

An example of the use of the heating effects of ultrasound is the treatment of sports injuries in muscles and joints (Figure 10.6.4). Intensity values of less than 30 000 W m^{-2} must be used in order to avoid tissue damage. Ultrasound exposures of 5 to 10 minutes are often part of a physiotherapy program. It is believed that the heating effect increases metabolism in the treated site and accelerates healing. This is most effective in bone and denser muscles, which are highly absorbing of sound waves. Mild heating by ultrasound can also be used to treat blockages of the middle ear region. Ultrasound is also being investigated as a treatment for arthritis.

FIGURE 10.6.4 High-intensity ultrasound is used to treat some sports injuries.

Destructive effects of ultrasound

If the intensity of the sound waves that are sent into the body is high enough, they can be used to destroy certain cells. The intense vibrations cause overheating and large stresses, which can rupture cell membranes. This property can be used beneficially, for example in the treatment of gallstones (in the gall bladder) or kidney stones. High-intensity ($\sim 10^5\,\mathrm{W\,m^{-2}}$) ultrasound waves physically break up the stones, and their component particles are then washed away by the normal removal processes of the body. Previously the removal of these items would have involved surgery.

Some types of tumorous cells can also be treated with ultrasound. Again, an intense beam of sound waves is used with the intention of breaking the cells apart. When this process is applied to a very concentrated area, surgeons have a very effective cauterising (cutting and sealing) tool. Neurosurgeons use extremely narrow beams of sound waves with an intensity of around $2.5 \times 10^5\,\mathrm{W\,m^{-2}}$ to 'cut out' brain tumours.

ACOUSTIC ENGINEERING—THEATRES AND SOUND SYSTEMS

Acoustic engineering uses the properties of sound waves to modify or improve people's listening experiences in a variety of venues. You will notice that in outdoor venues the acoustics are poor because the sound intensity gradually decreases further away from the speakers. In closed entertainment venues, such as a concert hall (Figure 10.6.5), there are several ways to take advantage of the properties of waves.

FIGURE 10.6.5 The concert hall of the Sydney Opera House.

Sound can be reflected, thus containing the sound energy, as discussed in Section 10.3.

Reflection from a flat, hard wall reflects and changes the direction of the sound. However, using only a flat, hard wall can have disadvantages in that there may be **echoes** or **reverberations**. In addition, standing waves may be formed in the room. For example, a subwoofer in a closed room creates regions of high-pressure variation or antinodes at the walls, thus creating a high-volume bass sound at the walls.

One way to reduce the volume of the bass at the walls is to use 'bass traps', which are usually made of foam and absorb the sound. In the home, foam or soft furnishings can absorb some of the bass sounds.

In a large-scale entertainment venue such as a concert hall, the goal is to contain as much sound energy as possible, so instead of using absorbers, modern theatres use diffusers. Diffusers reflect and disperse the sound in multiple directions (seen previously in Figure 10.3.7 on page 350). They have a non-flat, corrugated surface that is hard and acoustically non-absorbent (Figure 10.6.6).

FIGURE 10.6.6. An acoustic diffuser made of tinder blocks used in a recording studio.

A study on concert halls showed that flat surfaces also have the effect of de-emphasising and emphasising some frequencies so they do not faithfully reproduce the original sound. A diffuser, on the other hand, tends to reflect all frequencies and gives more of the original sound. At the same time, it reduces echoes and standing waves. In the home, the use of multiple reflecting surfaces, such as book cases, can act as a diffuser.

The design and placement of the seating (and the humans in them) in a theatre can cause diffraction effects, allowing more of the high-frequency sounds to be spread around the theatre. In the Helenistic Theatre of Epidaurus, an ancient theatre in Greece still used today, the corrugated effect of the seats acts to diffract sounds greater than 530 Hz, allowing a more faithful reproduction of sound throughout the theatre.

ACCOUSTIC ENGINEERING—NOISE POLLUTION

The properties of waves can also be used to minimise noise pollution from unwanted sources, such as roads. For example, housing estates built near major highways usually have high barriers with dense insulating material to reflect unwanted noise. In addition, houses near the roads are usually built with sensitive areas such as bedrooms further away from the road, and areas such as bathrooms, laundries and carports closer to the road. This allows for a large number of reflecting surfaces. In addition, vegetation, such as trees and dense shrubs on the side of the road, acts as a diffuser and absorber of sound.

Sound proofing a room requires several approaches. The use of double-glazed windows limits the propagation of sound from the outside, due to the small air gap between the two panes of glass. In addition, sound absorbers such as foam can be used for insulation. Structures on the walls of recording studios can be used to also reduce unwanted reflections (Figure 10.6.7).

FIGURE 10.6.7 In a recording studio foam is used to absorb sound and corrugations are used to reduce unwanted reflections.

10.6 Review

SUMMARY

- The wave intensity will decrease with the square of the distance from the source according to the relationship $I \propto \frac{1}{r^2}$.
- Ultrasound frequencies are above 20 000 Hz.
- Ultrasound can be used for medical imaging.
- High-intensity ultrasound can be used for heat treatment and for removing unwanted cells such as cancers and kidney stones.
- Acoustic engineering such as in theatres uses the properties of waves to optimise listening experiences.
- Noise pollution can be reduced by utilising the properties of reflection and understanding sound propagation.

KEY QUESTIONS

1 Earthquake waves are detected 250 km from the source with a strength of 2.5×10^4 W m^{-2}. What will the intensity of these waves be when they reach a major city at a distance of 1000 km from the source? Assume there is no loss from absorption between the source and where the waves are detected.

2 An ambulance started its siren 20 m after leaving the scene of an accident. After a short time, the intensity of the sound was measured as being $\frac{1}{8}$ of the original. Assuming the volume of the siren hadn't changed, how far away was the ambulance when the intensity was measured at the second position?

3 'High-intensity' sound waves are used in treating sports injuries. What does this mean?

4 An office has a lot of reverberation and echoes. How could this be reduced?

5 How does a diffuser reduce echoes and standing waves in a room without losing energy?

6 List two ways to reduce sound pollution from a busy road, and explain the physics of them.

Chapter review

10

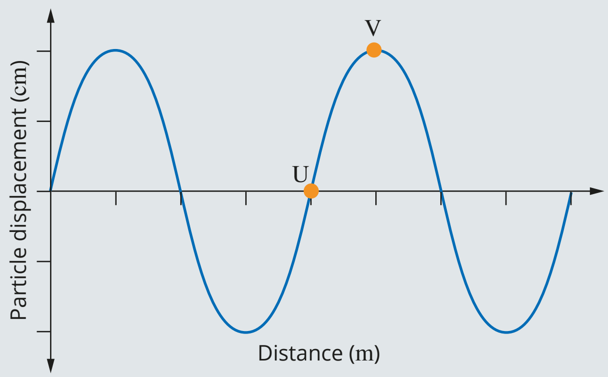
KEY TERMS

absorb
air pressure
amplitude
angle of incidence
angle of reflection
angle of refraction
antinode
compression
constructive interference
crest
critical angle
destructive interference
diffraction
diffuse
doppler effect
echo
energy
first harmonic
forcing frequency
frequency
fundamental
harmonic
intensity
inverse square law
longitudinal
mechanical wave
medium
natural frequency
node
normal
oscillate
overtone
particle displacement
period
phase
plane wave
pulse
rarefaction
ray
reflect
refraction
resonance
resonant frequency
reverberation
seismic wave
sinusoidal
standing wave
superposition
total internal reflection
transmit
transverse
travelling wave
trough
vibration
wave front
wavelength

1 Imagine that you watch from above as a stone is dropped into water. Describe the movement of the particles on the surface of the water.

2 Describe the similarities and differences between transverse and longitudinal waves.

3 At the moment in time shown on the graph, in what directions are the particles U and V moving?



4 The source of waves in a ripple tank vibrates at a frequency of 10.0 Hz. If the wave crests formed are 30.0 mm apart, what is the speed of the waves (in $m\,s^{-1}$) in the tank?

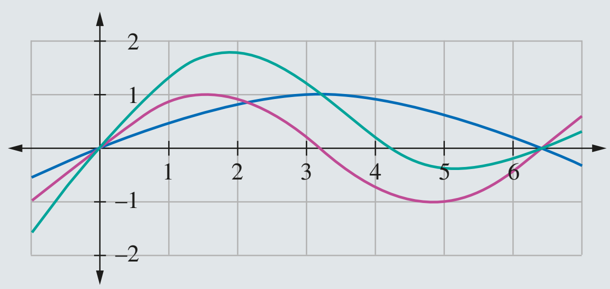
5 A submarine's sonar sends out a signal with a frequency of 32 kHz. If the wave travels at $1400\,m\,s^{-1}$ in seawater, calculate the wavelength of the signal.

6 A motor bike is able to produce a long, steady sound. You are unable to see the motor bike, but can hear the sound from it rise in frequency and then fall. Which one or more of the following options best explains the motion of the motor bike relative to you?

A The bike travelled towards you.
B The bike travelled away from you.
C The bike travelled past you.
D The bike travelled towards you, then away from you.

7 If you decreased the wavelength of the sound made by a loudspeaker, what effect would this have on the frequency and the velocity of the sound waves?

8 The following graph shows three wave forms. Two of the wave forms superimpose to form the third wave form:



Which wave is the result of the superposition of the other two?

9 Using ideas about the movement of particles in air, explain how you know sound waves only carry energy and not matter from one place to another.

10 A sound wave is emitted from a speaker and heard by Lee who is 50 m from the speaker. He made several statements once he heard the sound. Which one or more of the following statements made by Lee would be correct? Explain your answers.

   A Hearing a sound wave tells me that air particles have travelled from the speaker to me.
   B Air particles carried energy with them as they travelled from the speaker to me.
   C Energy has been transferred from the speaker to me.
   D Energy has been transferred from the speaker to me by the oscillation of air particles.

11 Describe the concept of resonance and why it would need to be considered when designing structures like buildings or bridges.

The following information relates to questions 11–13.

A signal generator is attached to a string vibrator producing vibrations down a string fixed at one end and free to move at the other. The string is kept at constant tension. The effective length of the string is 85 cm. The speed of the vibrations along the string is 340 m s^{-1}.

12 What is the lowest frequency of vibration that will produce a standing wave in the string?

13 What is the frequency of vibration of the third harmonic?

14 An earthquake causes a footbridge to oscillate up and down with a fundamental frequency once every 4.0 s. The motion of the footbridge can be considered to be like that of a string fixed at both ends. What is the frequency of the second harmonic for this footbridge?

15 The velocity of waves in a particular string at constant tension is 78 m s^{-1}. The string is fixed at both ends. If a particular frequency of a standing wave formed in the string is 428 Hz, how far apart would two adjacent antinodes be?

16 Reflection is possible from which of the following shaped surfaces if each surface is reflective? (More than one answer may be correct.)

   A flat surface
   B concave (curved in) surface
   C uneven surface
   D convex (curved out) surface

17 Resonance occurs when the frequency of a forcing vibration exactly equals the natural frequency of vibration of an object. Two special effects occur. Which one of the following responses relating to the effects of resonance is true? Explain your answer.

   A The amplitude of vibration will decrease.
   B The amplitude of vibration will increase.
   C The frequency of vibration will increase.
   D The frequency of vibration will decrease.

18 The third harmonic of an open pipe produces a frequency of 400 Hz. Assuming the speed of sound is 330 m s^{-1} calculate the length of the pipe?

19 A signal generator connected to a speaker produces sound waves that are directed into a tube closed at one end. The effective length of the tube is 85 cm and the speed of sound is 340 m s^{-1}.

   a What is the lowest frequency of sound that will produce resonance in the tube?
   b What frequency of sound will cause the tube to resonate at its third harmonic?

20 If a wave travels through water into air at an angle, what would you expect to happen to the angle of refraction? Explain your answer.

21 Under what conditions does total internal reflection occur?

22 The intensity of a light measured at some distance is x W m^{-2}. If the distance to the light was doubled what is the relative intensity of the light at this new distance?

   A $2x$
   B $4x$
   C $\frac{x}{2}$
   D $\frac{x}{4}$

23 Electromagnetic radiation also obeys the inverse square law. A radio transmitter produces a signal with an intensity of 1 W m^{-2} when measured at 1000 m. This signal was measured using a dish of 1 m^2. What diameter would a circular radio dish need to be to collect 1 W of signal at a distance of 4000 m? Give your answer to the nearest millimetre.

CHAPTER 11

Practical investigation

This chapter covers most of the skills needed to successfully plan and conduct a practical investigation.

Section 11.1 is a guide to designing and planning an investigation, including how to write a hypothesis, and how to identify the variables. It explains validity, reliability and accuracy, to assist in planning an investigation appropriately.

Section 11.2 is a guide to conducting investigations. It describes methods for accurately collecting and recording data to reduce errors. It explores presenting data using tables and graphs, to aid in selecting the most appropriate format for presenting the results.

Section 11.3 explains how to discuss an investigation and draw evidence-based conclusions that relate to the hypothesis and research question.

Practical investigation steps

The size and scope of a practical investigation can be initially quite daunting, but establishing a task list and timeline will help break it down into manageable steps. The entire task is expected to take between 7 and 10 hours.

Here are some steps that will need to be considered in a timeline:

- Determine the topic and type of investigation.
- Research and write down the theory on which the investigation is based.
- Determine an appropriate question to answer, and formulate a hypothesis.
- Identify the independent, dependent and controlled variables.
- Select equipment and resources needed for the investigation.
- Determine an appropriate procedure (methodology), considering validity, reliability and accuracy.
- Assess the risks and ethical issues and identify measures to address these.
- Conduct the investigation and record all data obtained.
- Analyse and evaluate the data.
- Evaluate your methods. Suggest ways of improving or extending the investigation.
- Write an evidence-based conclusion. Describe the limitations of the study.
- Write the final report or poster. (This should not be the focus of the investigation but rather an opportunity to communicate the investigation process and your conclusions.)

Some of these tasks are larger and will require more time than others. Many will overlap. Plan a realistic approach, consult with teachers to establish school-based time constraints and fix dates for the completion of each task. Allow time for reflection and to review your earlier work.

Key knowledge statements

Science Inquiry Skills

- identify, research, construct and refine questions for investigation; propose hypotheses; and predict possible outcomes.
- design investigations, including the procedure(s) to be followed, the materials required, and the type and amount of primary and/or secondary data to be collected; conduct risk assessments; and consider research ethics.
- conduct investigations, including using temperature, current and potential difference measuring devices, safely, competently and methodically for the collection of valid and reliable data.
- represent data in meaningful and useful ways, including using appropriate Système Internationale (SI) units and symbols, and significant figures; organise and analyse data to identify trends, patterns and relationships; identify sources of random and systematic error and estimate their effect on measurement results; identify anomalous data and calculate the measurement discrepancy between experimental results and a currently accepted value, expressed as a percentage; and select, synthesise and use evidence to make and justify conclusions.
- interpret a range of scientific and media texts, and evaluate processes, claims and conclusions by considering the quality of available evidence; and use reasoning to construct scientific arguments.
- select, construct, and use appropriate representations, including text and graphic representations of empirical and theoretical relationships, flow diagrams, nuclear equations, and circuit diagrams, to communicate conceptual understanding, solve problems and make predictions.
- select, construct, and use appropriate representations, including text and graphic representations of empirical and theoretical relationships, vector diagrams, free body/force diagrams, wave diagrams and ray diagrams, to communicate conceptual understanding, solve problems and make predictions.
- select, use, and interpret appropriate mathematical representations, including linear and non-linear graphs and algebraic relationships representing physical systems, to solve problems and make predictions.
- communicate to specific audiences and for specific purposes using appropriate language, nomenclature, genres, and modes, including scientific reports.

11.1 Designing and planning the investigation

FIGURE 11.1.1 There are many elements to a practical investigation, which may appear overwhelming to begin with. Taking a step-by-step approach will help the process and assist in completing a worthwhile investigation.

Taking the time to carefully plan and design a practical investigation before you begin will help you to maintain a clear and concise focus throughout. Preparation is essential. Ensure you understand the theory behind the investigation and prepare a detailed plan for the practical components of the investigation. This section is a guide to some of the key steps that should be taken when planning and designing a practical investigation (Figure 11.1.1).

DETERMINING THE TYPE OF INVESTIGATION

In your course, you will be required to undertake an investigation related to an area you have studied. Your teacher might suggest a particular topic or you may be required to suggest a topic that interests you. While your teacher may suggest the topic (such as heating processes) in most cases you will need to come up with the research question, hypothesis, variables, method, analysis and conclusion yourself.

Several types of research methods are used in science. Physics investigations typically fall into two methodologies:

- analysing the slope of a linear graph
- determining the relationship between two continuous **variables**.

From your research on your topic you may find that there is a mathematical relationship between the variables that you are exploring. It may be possible to use your algebra skills to write the mathematical relationship in the form of a linear equation, $y = mx + c$, where y is the dependent variable and x is the independent variable. A graph of the y and x variables will result in a linear graph with a slope equal to a constant, m, or collection of constants. By equating the value of the slope to these constants some meaningful analysis can be made, such as finding a value for the acceleration due to gravity or the specific heat capacity of a metal.

If, in your research, you find that there is no equation that links your variables, then you can investigate to determine what the relationship might be. By graphing your variables, you may find that the relationship is linear, or it may be inverse, exponential or even inverse squared. An investigation of this kind results in a calibration curve, which can be used to predict values of the dependent variable given particular values of the independent variable.

DEVELOPING RESEARCH QUESTIONS AND AIMS, FORMULATING HYPOTHESES AND IDENTIFYING VARIABLES

The research question, aim and hypothesis are interlinked. It is important to note that each of these can be refined as the planning of the investigation continues.

Formulating a question

A research question is a question that comes from an inquiring mind. When you are actively involved in developing your scientific understanding then you will want to know what factors affect a variable, or what is the relationship between two variables. The research question poses the question that the investigation seeks to answer. For example: What is the relationship between voltage and current?

Before formulating a question, it is good practice to conduct a literature review of the topic to be investigated. You should become familiar with the relevant scientific concepts and key terms.

During this review, write down questions, correlations, or equations as they arise.

Compile a list of possible ideas. Do not reject ideas that initially might seem impossible. Use these ideas to generate questions that are answerable.

Before constructing a hypothesis, formulate a question that needs an answer. This question will lead to a hypothesis when:

- the question is reduced to measurable variables
- a prediction is made based on knowledge and experience.

Evaluating your question

Once a question has been chosen, stop to evaluate the question before progressing. The question may need further refinement or even further investigation before it is suitable as a basis for an achievable and worthwhile investigation. A major planning point is to attempt something that it is possible to complete in the time available or with the resources on hand. It might be a little difficult to create a particularly complicated device with the facilities available in the school laboratory.

To evaluate the question, consider the following:

- Relevance: Is the question related to the appropriate area of study?
- Clarity and measurability: Can the question lead to a clear hypothesis? If the question cannot lead to a specific hypothesis, then it is going to be very difficult to complete the research.
- Time frame: Can the question be answered within a reasonable period of time? Is the question too broad?
- Knowledge and skills: Do you have a level of knowledge and a level of laboratory skills that will allow the question to be explored? Keep the question simple and achievable.
- Practicality: Are resources, such as laboratory equipment and materials, likely to be readily available? Keep things simple. Avoid investigations that require sophisticated or rare equipment. More-readily-available equipment includes timing devices, objects that could be used as projectiles, a tape measure and other common laboratory equipment.
- Safety and ethics: Consider the safety and ethical issues associated with the question you will be investigating. If there are issues, can these be addressed?
- Advice: Seek advice from the teacher about the question. Their input may prove very useful. Their experience may lead them to consider aspects of the question that you have not thought about.

Defining the aim of the investigation

An aim is a statement describing in detail what will be investigated to answer the research question. For example: The aim of the experiment is to investigate the relationship between the voltage and current in a circuit of constant resistance. Each aim should directly relate to the variables that will be referred to in the hypothesis. The aims do not need to include the details of the method.

Example

- Aim: The aim of the experiment is to investigate the relationship between mass and acceleration, when a constant force is applied.

Hypothesis

A hypothesis is a definite statement, based on previous knowledge and evidence or observations, that attempts to answer the research question. The hypothesis must relate the independent and dependent variables and describe the relationship between them. For example: Increasing the voltage supplied to a circuit of constant resistance increases the current in a linear way.

Here are some examples of hypotheses:

- For a constant force, if the mass is increased, the acceleration is decreased as an inverse relationship.
- If the value of the resistance of a circuit increases, the current flowing in the circuit will decrease as an inverse relationship.

- Assuming no heat loss to the surroundings, the temperature rise of a fixed mass of water is proportional to the time it is heated by a constant power source.
- As the height an object is dropped from increases, the final velocity of the object will increase as a squared relationship.

There are no wrong or right hypotheses. You might formulate a hypothesis that a more experienced person will disagree with; however, the purpose of an investigation is to find the answer to a research question. If the answer to the question supports your hypothesis, then that is a positive result, as it will confirm your understanding of the concept. On the other hand, if your investigation does not support your hypothesis, then that is a positive result as well, as you can now say that your original understanding was not correct and you can change your understanding to a more scientific one. Some of you might notice that the following hypothesis will not be supported by the investigation:

- The greater the mass of a marble, the faster it will hit the ground, when dropped from the same height.

This doesn't mean that the hypothesis is wrong, but it may indicate that there was some misconception that you had that was not exposed in your literature review.

Formulating a hypothesis

A good hypothesis should:

- be a definite statement of the relationship
- include an independent and a dependent variable that is continuous and measurable
- be worded so that it can be tested in the experiment.

The hypothesis should also be falsifiable. This means that a negative outcome would disprove it. For example, the hypothesis that all apples are round cannot be proved beyond doubt, but it can be disproved—in other words, it is falsifiable. In fact, only one oval-shaped apple is needed to disprove this hypothesis. Unfalsifiable hypotheses cannot be proved by science. These include hypotheses on ethical, moral and other subjective judgements.

Variables

A good scientific hypothesis can be tested—that is, it can be supported or refuted through investigation. To be a testable hypothesis, it should be possible to measure both the change or treatment and the effect, or what will happen. The factors that can be changed, or are changed as a result of the experiment or investigation, are called the variables. An experiment or investigation determines the relationship between variables.

There are three categories of variables:

- The **independent variable** is the variable that is controlled by the researcher (the one that is selected and changed). You must test only one independent variable in any investigation, otherwise it cannot be stated that the changes in the dependent variable are the result of changes in the independent variable.
- The **dependent variable** is the variable that may change in response to a change in the independent variable. This is the variable that will be measured or observed. You should measure only one dependent variable in any investigation. If you want to measure another dependent variable then you will need to do another investigation with another hypothesis.
- **Controlled variables** are all the variables that must be kept constant during the investigation otherwise the test cannot be fair.

  Read the following example of a hypothesis.

  If the cross-sectional area of a resistor is constant, the longer the wire, the greater the resistance as a linear relationship.

Identify the different variables.

- independent variable: length of wire
- dependent variable: resistance of the wire
- controlled variables: potential difference, material of the resistor, temperature of the resistor.

Completing a table like Table 11.1 will assist in evaluating the question or questions.

TABLE 11.1.1 Break the question down to determine the variables.

Research question	How does the power of a kettle affect the time taken to boil water?
Independent variable	the power of the kettle
Dependent variable	the time the kettle takes to boil water
Controlled variables	mass of the water, purity of the water, starting temperature of the water and kettle
Potential hypothesis	The greater the power of a kettle, the less time it will take to boil water as an inverse relationship

Qualitative and quantitative variables

Variables are either qualitative or quantitative, with further subsets in each category.

- **Qualitative variables** can be observed but not measured; for example, describing a light globe as bright or dim. They can only be sorted into groups or categories such as brightness, type of material of construction or type of device.
  - Nominal variables are categorical variables in which the order is not important; for example, the type of material or type of device.
  - Ordinal variables are categorical variables in which order is important and groups have an obvious ranking or level; for example, brightness (Figure 11.1.2).
- **Quantitative variables** can be measured. Length, area, weight, temperature and cost are all examples of quantitative data.
  - Discrete variables consist of only integer numerical values, not fractions; for example, the number of pins in a packet, the number of springs connected together, or the energy levels in atoms.
  - Continuous variables allow for any numerical value within a given range; for example, the measurement of temperature, length, weight, and frequency.

In Physics, you should ensure that you choose quantitative variables for both the independent and dependent variables. This will allow you to construct a line graph, and therefore determine the slope of the line, or the relationship between the variables.

FIGURE 11.1.2 When recording qualitative data, describe in detail how each variable will be defined. For example, if recording the brightness of light globes, light meters are a quantitative way to gather data.

WRITING THE METHODOLOGY

The methodology, or method, of your investigation is a step-by-step procedure. When detailing the method, ensure it enables you to conduct a valid, reliable and accurate investigation.

Validity

Validity refers to whether an experiment is in fact testing the hypothesis. Is the investigation obtaining data that is relevant to the question, or is it flawed?

To ensure an investigation is valid, it should be designed so that only one variable is being changed at a time. The remaining variables must remain constant, so that meaningful conclusions can be drawn about the effect of the independent variable alone.

To ensure validity, you must carefully determine:

- the independent variable—the variable that will be changed, and how it will change
- the dependent variable—the variable that will be measured
- the controlled variables—the variables that must remain constant.

Reliability

Reliability refers to the idea that the experiment can be repeated many times and will obtain consistent results. You can maintain the investigation's reliability by:

- listing and defining the control variables and how they will be kept constant
- listing the detailed steps that you will take to conduct the experiment, describing what you will do and how you will measure and record data
- ensuring that there are enough changes of the independent variable. Typically, five changes over a wide range of the independent variable are considered sufficient.
- ensuring there are enough trials conducted for each value of the independent variable. Typically, you should conduct at least three trials repeating the experiment, then average the three measurements. This reduces random errors and allows systematic errors to be identified. If a reading differs too much from the rest (known as an outlier), discard it before averaging (Figure 11.1.3).

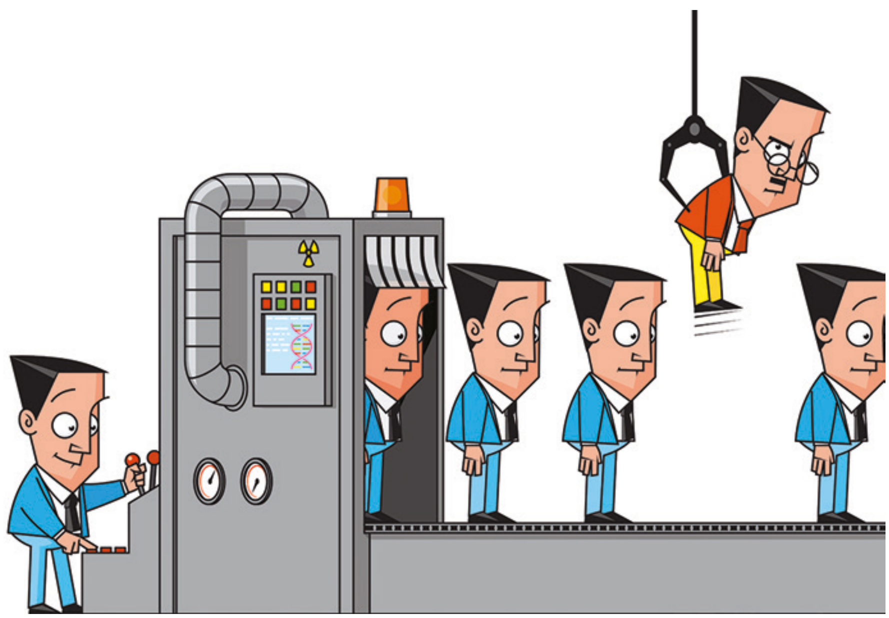


FIGURE 11.1.3 Replication increases the reliability of your investigation. It ensures that if anyone repeats the investigation they will obtain similar data.

Accuracy and precision

Precision refers to the extent to which the instrument can make repeated measures of the dependent variable that are the same for the same value of the independent variable. For example, if each measurement of the current in an electrical circuit is within 0.1 A of the others, then the device is more precise than a device in which there is a difference of 0.5 A. Accuracy refers to the ability to obtain the correct measurement.

You will need to consider if the instruments to be used are sensitive enough. Build some testing into your investigation to confirm the accuracy and reliability of the equipment and your ability to read the information obtained.

Reasonable steps to ensure the accuracy of the investigation include considering:

- the type of instrument that will be used to measure the independent and dependent variables.
- calibrating the measuring equipment by testing a standard.

Describe the materials and method in appropriate detail in your logbook. This should ensure that every measurement can be repeated and the same result obtained within reasonable margins of experimental error (less than 5% is reasonable).

Data analysis

Data analysis is part of the method. Consider how the data will be presented and analysed. A wide range of analysis tools are available. For example, tables can be used to organise data so that patterns can be established, and graphs can show relationships and comparisons. In fact, preparing an empty table showing the data that needs to be obtained will help in the planning of the investigation. See pXXX for more information on organizing a data table.

Sourcing appropriate materials and technology

When designing your investigation, you will need to decide on the materials, technology and instrumentation that will be used to carry out your research. It is important to find the right balance between items that are easily accessible and those that will give you accurate results. As you move onto conducting your investigation, it will be important to take note of the precision of your chosen instruments and how this affects the accuracy and validity of your results.

Modifying the methodology

The methodology may need modifying as the investigation is carried out. The following actions will help to determine any issues in the methodology and how to modify them:

- Record everything.
- Be prepared to make changes to the approach.
- Note any difficulties encountered and the ways they were overcome. What were the failures and successes? Every test carried out can contribute to the understanding of the investigation as a whole, no matter how much of a disaster it may first appear.
- Do not panic. Go over the theory again, and talk to the teacher and other students. A different perspective can lead to a solution.

If the expected data is not obtained, don't worry. As long as it can be critically and objectively evaluated, the limitations of the investigation are identified and further investigations proposed, the work is worthwhile.

COMPLYING WITH ETHICAL AND SAFETY GUIDELINES

Ethical considerations

Some investigations require an ethics approval—consult with the teacher. In fact, when deciding on an investigation, identify all possible ethical considerations and evaluate whether those parts of the investigation are necessary or if there are ways you can reduce or mitigate them.

Occupational health and safety

While planning for an investigation, it is important for the safety of yourself and the safety of others that the potential risks are considered.

FIGURE 11.1.4 When planning an investigation, you need to identify, assess and control hazards.

Everything we do has some risk involved. Risk assessments are performed to identify, assess and control hazards. A risk assessment should be performed for any situation, in the laboratory or outside in the field. Always identify the risks and control them to keep everyone safe. For example, carry out voltage–current experiments with low voltages (less than 6.0 V DC or 4 × 1.5 V batteries) coupled to resistors so that the currents in the circuits are of the order of milliamps. *At all times* avoid direct exposure to 240 V AC household voltages (Figure 11.1.4).

To identify risks, think about:

- the activity that will be carried out
- the equipment or chemicals that will be used.

The following hierarchy of risk controls is organised from most effective to least effective:

1 *Elimination:* Eliminate dangerous equipment, procedures or substances.
2 *Substitution:* Find different equipment, procedures or substances to use that will achieve the same result, but have less risk associated with them.
3 *Isolation:* Ensure there is a barrier between the person and the hazard. Examples include physical barriers such as guards in machines, or fume hoods to work with volatile substances.
4 *Engineering controls:* Modify equipment to reduce risks.
5 *Administrative controls:* Provide guidelines, special procedures, warning signs and safe behaviours for any participants.
6 *Personal protective equipment (PPE):* Wear safety glasses, lab coats, gloves, and respirators etc. where appropriate, and provide these to other participants.

Science outdoors

Sometimes investigations and experiments will be carried out outdoors. Working outdoors has its own set of potential risks and it is equally important to consider ways of eliminating or reducing these risks.

As an example, Table 11.1.2 contains examples of risks associated with field work outdoors.

TABLE 11.1.2 Risks associated with fieldwork outdoors.

Risks	Control measures
sunburn	wear sunscreen, a hat and sunglasses
hot or cold weather	wear clothing to protect against heat or cold
projectile launch	create barriers so that people know not to enter the area
trip hazards	minimise the use of cables (electrical, computer) and cover them up with matting be aware of tree roots, rocks etc.

First aid measures

Minimising the risk of injury reduces the chance of requiring first aid assistance. However, it is still important to have someone with first aid training present during practical investigations. Always tell the teacher or laboratory technician if an injury or accident happens.

Personal protective equipment

Everyone who works in a laboratory wears items that help keep them safe. This is called **personal protective equipment (PPE)** and includes:

- safety glasses
- shoes with covered tops
- disposable gloves when handling chemicals
- a disposable apron or a lab coat if there is risk of damage to clothing
- ear protection if there is risk to hearing.

11.1 Review

SUMMARY

- An aim is a statement describing in detail what will be investigated. For example: The aim of the experiment is to investigate the relationship between force, mass and acceleration.
- A hypothesis is a definite statement of the relationship between the independent and dependent variables based on previous knowledge and evidence or observations that attempts to answer the research question. For example: With the force kept constant, the acceleration decreases with increasing mass as an inverse relationship.
- Once a question has been chosen, stop to evaluate the question before progressing. The question may need further refinement or even further investigation before it is suitable as a basis for an achievable and worthwhile investigation. Make sure that it is possible to complete the activity in the time available and with the resources on hand. It might be a little difficult to create a particularly complicated device with the facilities available in the school laboratory.
- There are three categories of variables:
  - The independent variable is the variable that is controlled by the researcher (the one that is selected and changed).
  - The dependent variable is the variable that may change in response to a change in the independent variable. This is the variable that will be measured or observed.
  - Controlled variables are all the variables that must be kept constant during the investigation so that it is a fair test.
- The methodology of your investigation is a step-by-step procedure. When detailing the methodology, ensure it complies as a valid, reliable and accurate investigation.
- It is also important to determine how many times the independent variable needs to be changed and how many trials need to be run for each change in the independent variable.
- Data analysis is part of the method. Consider how the data will be presented and analysed. A wide range of analysis tools could be used. For example, tables organise data so that patterns can be established and graphs can show relationships and comparisons.
- In every investigation you need to consider the risks and potentially hazardous situations and act to minimise those risks.

KEY QUESTIONS

1 In a practical investigation the student changes the voltage by adding or subtracting batteries in series to the circuit.
  a How could the voltage be a discrete value?
  b How could it be continuous?

2 In another experiment the student uses the following range of values to describe the brightness of a light: dazzling, bright, glowing, dim, off
What type of measurement is the variable 'brightness'?

3 Select the best hypothesis from the three options below. Give reasons for your choice.
  A Hypothesis 1: The greater the mass of the marble you drop, the greater the final velocity of the marble as a linear relationship.
  B Hypothesis 2: The greater the potential difference across a resistor the greater the current through it.
  C Hypothesis 3: Different metals will have different resistances.

4 Give the correct term that describes an experiment with each of the following conditions.
  a The experiment addresses the hypothesis and aims.
  b The experiment is repeated and consistent results are obtained.
  c Appropriate equipment is chosen for the desired measurements.

5 A student wanted to find out how the tension in an elastic band affects the initial velocity of it when launched from her finger. State:
  a the independent variable
  b the dependent variable
  c three controlled variables.

11.2 Conducting investigations and recording and presenting data

FIGURE 11.2.1 When carrying out your investigation try to maintain high standards to minimise potential errors.

Once the planning and design of a practical investigation is complete, the next step is to undertake the investigation and record the results. As with the planning stages, there are key steps and skills to keep in mind to maintain high standards and minimise potential error throughout the investigation (Figure 11.2.1).

This section will focus on the best methods of conducting a practical investigation, of systematically generating, recording and processing data and of presenting it in a concise and clear manner.

CONDUCTING INVESTIGATIONS TO COLLECT AND RECORD DATA

For an investigation to be scientific, it must be objective and systematic. Ensuring familiarity with the methodology and protocols before beginning will help you to achieve this.

When working, keep asking questions. Is the work biased in any way? If changes are made, how will they affect the study? Will the investigation still be valid for the aim and hypothesis?

It is essential that during the investigation the following are recorded in the logbook:

- all quantitative data collected
- the methods used to collect the data
- any incident, feature or unexpected event that may have affected the quality or validity of the data.

The data recorded in the logbook is the **raw data**. Usually this data needs to be processed in some manner before it can be presented. If an error occurs in the processing of the data or you decide to present the data in an alternative format, the recorded raw data will always be available for you to refer back to.

IDENTIFYING ERRORS

Most practical investigations have errors associated with them. Errors can occur for a variety of reasons. Being aware of potential errors helps you to avoid or minimise them. For an investigation to be accurate, it is important to identify and record any errors.

There are two types of errors:

- systematic errors
- random errors.

FIGURE 11.2.2 To avoid a systematic error, make sure that you are using measuring equipment correctly. Laser speed guns, for example, need to be placed on a stationary support so the aim point is held on a single target point for the duration of the read.

Systematic errors

A **systematic error** is an error that is consistent and will occur again if the investigation is repeated in the same way.

Systematic errors are usually a result of instruments that are not calibrated correctly or methods that are flawed.

An example of a systematic error would be if a ruler mark indicating 5 cm from 0 cm was actually only 4.9 cm from 0 cm due to a manufacturing error or shrinkage of the wood. Another example would be if the researcher repeatedly used a piece of equipment incorrectly throughout the entire investigation. Figure 11.2.2 shows how traffic police reduce systematic errors in their data collection.

Random errors

Random errors occur in an unpredictable manner and are generally small. A random error could be, for example, the result of a researcher reading the same result correctly one time and incorrectly another time. Another example would be if an instrument were affected by a power cut, or low battery power.

Techniques for reducing error

Designing the method carefully, including selection and use of equipment, will help reduce errors.

Appropriate equipment

Use the equipment that is best suited to the data that needs to be collected to validate the hypothesis. Determining the units of the data being collected and at what scale will help to select the correct equipment. Using the right unit and scale will ensure that measurements are more accurate and precise (with smaller systematic errors).

Significant figures are the numbers that convey meaning and precision. The number of significant figures used depends on the scale of the instrument. It is important to record data to the number of significant figures available from the equipment or observation. Using either a greater or smaller number of significant figures can be misleading.

Review the following examples to learn more about significant figures:

- 15 has two significant figures
- 3.5 has two significant figures
- 3.50 has three significant figures
- 0.037 has two significant figures
- 1401 has four significant figures.

To calculate gravitational potential energy (E_g), the formula is $E_g = mg\Delta h$.

If $g = 9.80\,\text{m}\,\text{s}^{-2}$, mass = 7.50 kg, height = 0.64 m (64 cm):

$E_g = 9.80 \times 7.50 \times 0.64 = 47.09\,\text{J}$

But only quote the answer to the least number of significant figures in the data; that is, to two significant figures, so $E_g = 47\,\text{J}$.

Although digital scales can measure to many more than two figures and calculators can give 12 figures, be sensible and follow the significant figure rules.

Calibrated equipment

Some equipment, such as some motion sensors, needs to be calibrated before use to account for the temperature at the time. Before carrying out the investigation, make sure the instruments or measuring devices are properly calibrated and are, in general, functioning correctly. For example, measure the temperature and apply a correction to the speed of sound to calibrate a motion sensor if necessary.

Correct use of equipment

Use the equipment properly. Ensure any necessary training has been done to use the equipment and that you have had an opportunity to practice using the equipment before beginning the investigation. Improper use of equipment can result in inaccurate, imprecise data with large errors, and the validity of the data can be compromised.

Incorrect reading of measurements is a common misuse of equipment. Make sure all the equipment needed in the investigation can be used correctly and record the instructions in detail so they can be referred back to if the data doesn't appear correct.

RECORDING AND PRESENTING QUANTITATIVE DATA

Raw data is unlikely to be used directly to validate the hypothesis. However, raw data is essential to the investigation and plans for collecting the raw data should be made carefully. Consider the formulas or graphs that will be used to analyse the data at the end of the investigation. This will help to determine the type of raw data that needs to be collected in order to validate the hypothesis.

For example, to calculate take-off velocity for a vertical jump, three sets of raw data will need to be collected using a force platform: the athlete's standing body weight, the ground reaction force and the time during the vertical jump. The data can then be processed to obtain the take-off impulse.

Once you have determined the data that needs to be collected, prepare a table in which to record the data.

ANALYSING AND PRESENTING DATA

The raw data that has been obtained needs to be presented in a way that is clear, concise and accurate.

There are a number of ways of presenting data, including tables, graphs, flow charts and diagrams. The best way of visualising the data depends on its nature. Try several formats before making a final decision, to create the best possible presentation.

Presenting raw and processed data in tables

Tables organise data into rows and columns and can vary in complexity according to the nature of the data. Tables can be used to organise raw data and processed data or to summarise results.

The simplest form of a table is a five-column format. In a five-column table, the first column should contain the independent variable (the one being changed) and the second to fourth columns should contain the three trials of the dependent variable (the one that may change in response to a change in the independent variable). The final column should contain the average of the three trials.

Tables should have the following features:

- a descriptive title that contains both the independent and depended variables
- column headings (including the unit and the uncertainty)
- aligned figures (align the decimal points)
- the independent variable placed in the left column
- the three trials of the dependent variable placed in the right columns with the average column on the end.

Processed data with the units and uncertainty if this is required for graphing e.g. if T^2 needs to be plotted instead of T.

Look at the table in Figure 11.2.3, which has been used to organise raw and processed data about the effect of current on voltage.

Effect of potential difference on the current through a resistor ← clear title

Potential difference (V)	Current trial 1 ± 0.01 (A)	Current trial 2 ± 0.01 (A)	Current trial 3 ± 0.01 (A)	Average current (A)
1.50	0.31	0.34	0.32	0.32 ± 0.02
2.00	0.42	0.45	0.41	0.43 ± 0.02
2.50	0.50	0.51	0.52	0.51 ± 0.01
3.00	0.62	0.65	0.58	0.62 ± 0.04
3.50	0.70	0.71	0.70	0.70 ± 0.01

headings for each column with units and uncertainties

consistent number of significant figures

independent variable

consistent number of decimal places

averages with uncertainties

FIGURE 11.2.3 A simple table listing the raw data obtained in the second, third and fourth columns and processed data in the fifth column.

A table of processed data usually presents the average values of trials, the **mean**. However, the mean on its own does not provide an accurate picture of the results.

To report processed data more accurately, the uncertainty should be presented as well.

Uncertainty

When there is a range of measurements of a particular value, the mean must be accompanied by the **uncertainty**, for your results to be presented as a mean in an accurate way. In other words, the mean must be accompanied by a description of the range of data obtained.

Uncertainty = ± (maximum variance from the mean)

For example, the speed, in $km\,h^{-1}$, of cars travelling down a certain road was:

46, 50, 55, 48, 50, 58, 45

The average speed would be:

$(46 + 50 + 55 + 48 + 50 + 58 + 45) \div 7 = 50\,km\,h^{-1}$

The uncertainty would be the maximum variance from the average: 58 is 8 above the average, so the uncertainty is 8.

This data should be presented as:

Average speed is $50 \pm 8\,km\,h^{-1}$.

Other descriptive statistics measures

The mean and the uncertainty are statistical measures that help describe data accurately. Other statistical measures that can be used, depending on the data obtained, are:

- **mode**: the mode is the value that appears most often in a data set. This measure is useful to describe qualitative or discrete data (for example, the mode of the values 0.01, 0.01, 0.02, 0.02, 0.02, 0.03, 0.04 is 0.02).
- **median**: the median is the 'middle' value of an ordered list of values (for example, the median of the values 5, 5, 8, 8, 9, 10, 20 is 8). The median is used when the data range is spread, for example, due to the presence of unusual results, making the mean unreliable.

Graphs

In general, tables provide more detailed data than graphs, but it is easier to observe trends and patterns in data in graphical form than in tabular form.

Graphs are used when two variables are being considered and one variable is dependent on the other. The graph shows the relationship between the variables.

There are several types of graphs that can be used, including line graphs, bar graphs and pie charts. The best one to use will depend on the nature of the data.

General rules to follow when making a graph (Figure 11.2.4) include the following:

- Keep the graph simple and uncluttered.
- Use a descriptive title that contains the independent and dependent variables.
- Represent the independent variable on the x-axis and the dependent variable on the y-axis.
- Make axes proportionate to the data.
- Clearly label axes with both the variable and the unit in which it is measured.
- Include error bars showing the uncertainty of each point. The error bar should extend above and below the plotted point equal to the uncertainty in the dependent variable and to the left and right of the plotted point equal to the uncertainty of the independent variable.

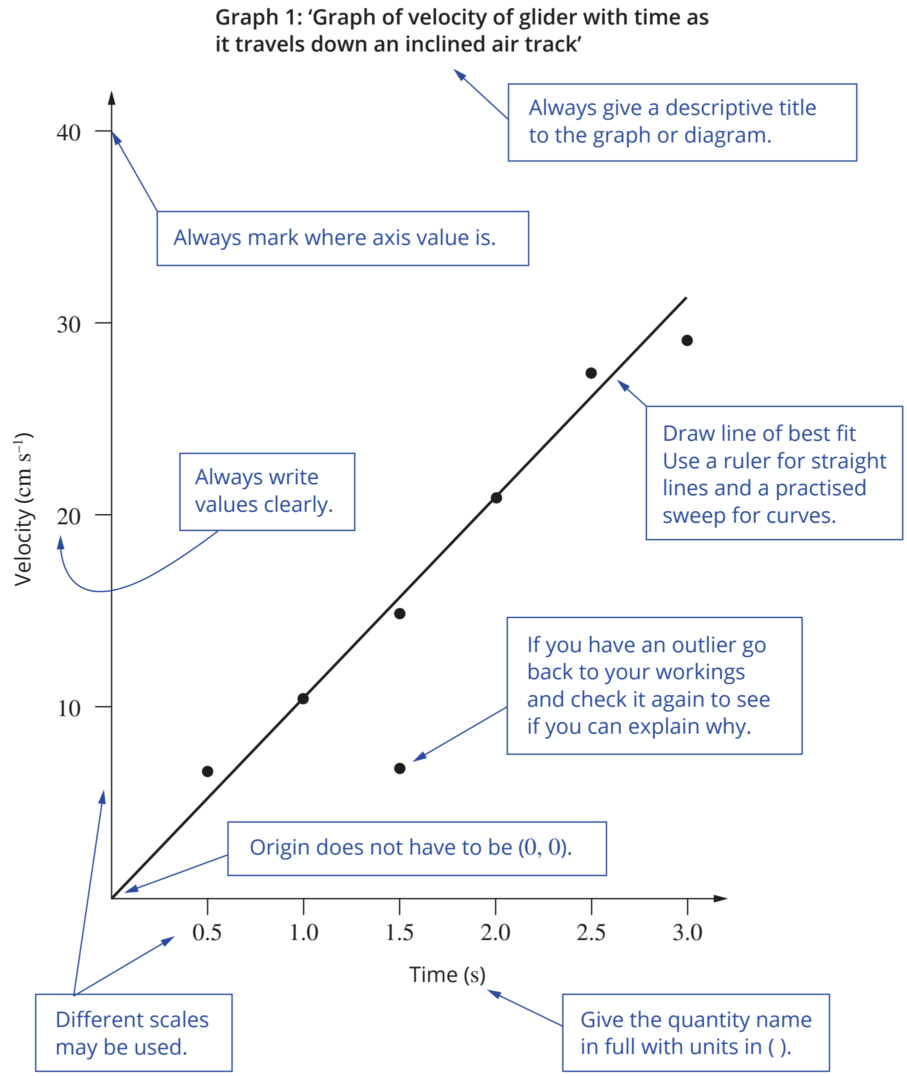


FIGURE 11.2.4 A graph shows the relationship between two variables.

Line graphs

Line graphs are a good way of representing continuous quantitative data. In a line graph, the values are plotted as a series of points with error bars on the graph. There are two ways of joining these points:

- A line can be ruled from each point to the next (Figure 11.2.5a). It shows the overall trend; it is not meant to predict the value of the points between the plotted data. These graphs are seldom used in Physics.
- The points can be joined with a single smooth straight or curved line (Figure 11.2.5b). This creates a trend line, also known as a line of best fit. The line of best fit does not have to pass through every point but should go through as many error bars as possible. It is used when there is an obvious trend between the variables. These graphs are most commonly used in Physics.

Outliers

Sometimes when the data is collected, there may be one point that does not fit with the trend and is clearly an error. This is called an **outlier**. An outlier is often caused by a mistake made in measuring or recording data, or from a random error in the measuring equipment. If there is an outlier, include it on the graph, but ignore it when adding a line of best fit (as in Figure 11.2.4, where the point (1.5, 6) is an outlier).

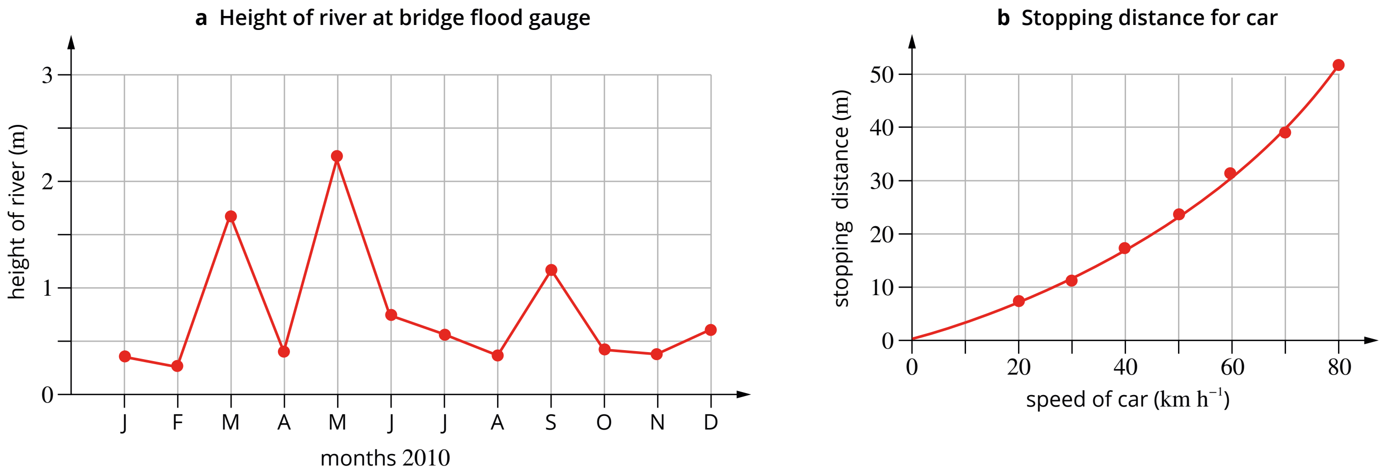


FIGURE 11.2.5 (a) The data in the graph is joined from point to point. (b) The data in graph is joined with a line of best fit, which shows the general trend.

11.2 Review

SUMMARY

- It is essential that during the investigation, the following are recorded in the logbook:
  - all quantitative data collected
  - the methods used to collect the data
  - any incident, feature or unexpected event that may have affected the quality or validity of the data.
- A systematic error is an error that is consistent and will occur again if the investigation is repeated in the same way. Systematic errors are usually a result of instruments that are not calibrated correctly or methods that are flawed.
- Random errors occur in an unpredictable manner and are generally small. A random error could be, for example, the result of a researcher reading the same result correctly one time and incorrectly another time.
- The number of significant figures used depends on the scale of the instrument used. It is important to record data to the number of significant figures available from the equipment or observation.
- The simplest form of a table is a five-column format in which the first column contains the independent variable (the one being changed), the second to fourth columns contain the trials of the dependent variable (the one that may change in response to a change in the independent variable) and the final column contains the average of the trials.
- When there is a range of measurements of a particular value, the average must be accompanied by the uncertainty.
- General rules to follow when making a graph include the following:
  - Keep the graph simple and uncluttered.
  - Use a descriptive title that contains the independent and dependent variables.
  - Represent the independent variable on the x-axis and the dependent variable on the y-axis.
  - Make axes proportionate to the data.
  - Clearly label axes with both the variable and the unit in which it is measured.
  - Include error bars showing the uncertainty of each point. The error bar should extend above and below the plotted point equal to the uncertainty in the dependent variable and to the left and right of the plotted point equal to the uncertainty of the independent variable.

KEY QUESTIONS

1 The masses of 1 cm^3 cubes of potato were recorded and the cubes placed in distilled water. After 60 minutes, the cubes were weighed again and the difference in mass was calculated. What type of error is involved:

a if the electronic scales only measured in 1 g increments?

b if the electronic scales were affected briefly by a power surge?

2 If using the quantities mass = 7.50 kg and speed = 1.4 $m\,s^{-1}$ in a calculation, what would be the appropriate number of significant figures in the answer?

3 For the data set 21, 28, 19, 19, 25, 24 determine:

a the mean

b the mode

c the median

d the uncertainty in the mean.

4 Plot the following data set with error bars, assigning each variable to the appropriate axis on the graph.

Current (A) ± 0.01 (A)	Voltage (V) ± 0.01 (V)
0.06	2.07
0.05	1.56
0.04	1.24
0.03	0.93
0.02	0.63

5 How can the general pattern (trend) of the data set in question 4 be represented once the points are plotted?

11.3 Discussing investigations and drawing evidence-based conclusions

Now that the chosen topic has been thoroughly researched and the investigation has been conducted and data collected, it is time to draw it all together. The final part of the investigation involves summarising the findings in an objective, clear and concise manner.

FIGURE 11.3.1 To discuss and conclude your investigation, use the raw and processed data.

EXPLAINING RESULTS IN THE DISCUSSION

The discussion is the part of the investigation where the evaluation and explanation of the investigation methods and results takes place. It is the interpretation of what the results mean.

The key sections of the discussion are:

- analysing and evaluating data
- evaluating the investigative method
- explaining the link between the investigation findings and the relevant physics concepts.

When writing the discussion, consider the message you want to convey to the audience. Statements should be clear and concise. At the conclusion of the discussion, the audience must have a clear idea of the context, results and implications of the investigation.

ANALYSING AND EVALUATING DATA

In the discussion, the findings of the investigation need to be analysed and interpreted.

- State whether a pattern, trend or relationship was observed between the independent and dependent variables. Describe what kind of pattern it was and specify under what conditions it was observed.
- Were there discrepancies, deviations or anomalies in the data? If so, these should be acknowledged and explained.
- Identify any limitations in the data you have collected. Perhaps a larger sample or further variations in the independent variable would lead to a stronger conclusion.

Trends in line graphs

Graphs are drawn to show the relationship, or trend, between two variables, as shown in Figure 11.3.2.

- Variables that change in linear or direct proportion to each other produce a straight, sloping trend line.
- Variables that change exponentially in proportion to each other produce a curved trend line.
- When there is an inverse relationship, one variable increases as the other variable decreases.
- When there is no relationship between two variables, one variable will not change even if the other changes.

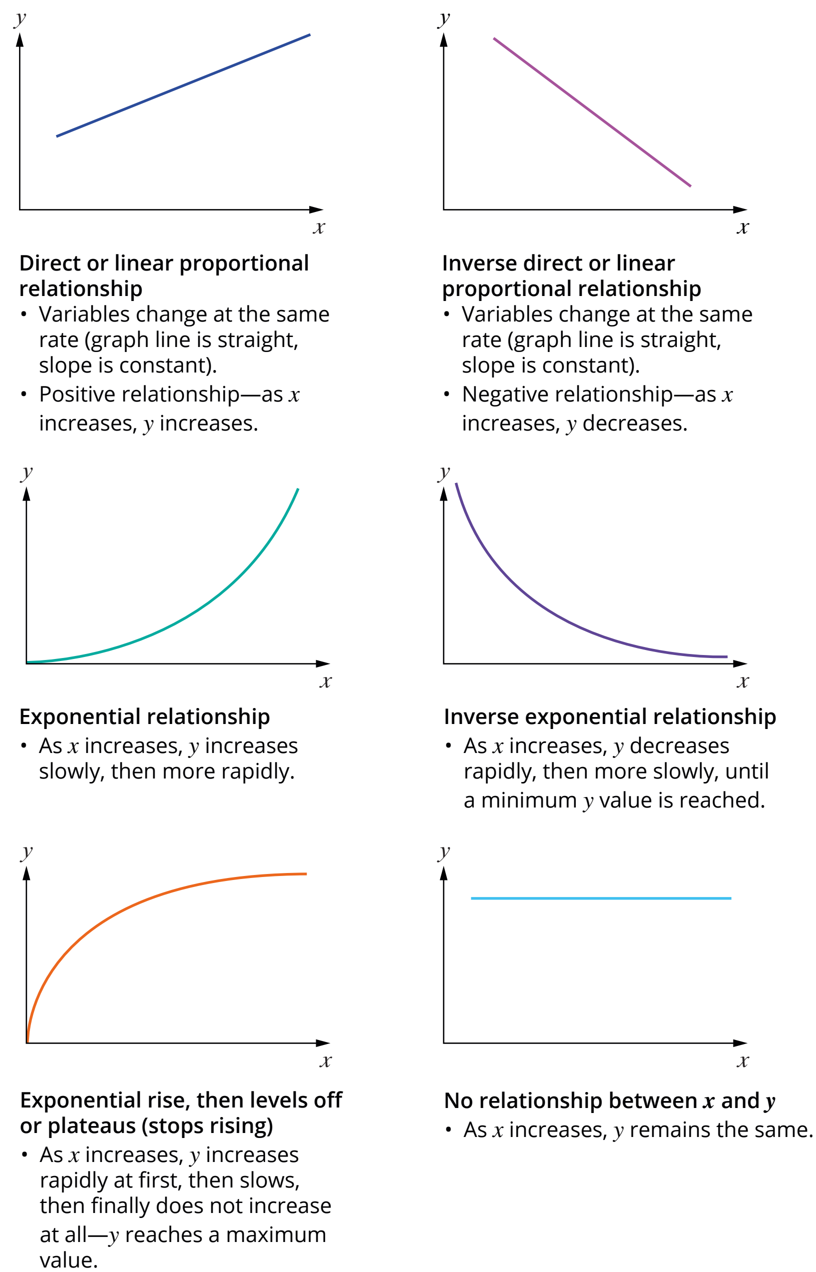


FIGURE 11.3.2 Various relationships can exist between two variables.

Remember that the results may be unexpected. This does not make the investigation a failure. However, the findings must be related to the hypothesis, aims and method.

EVALUATING THE METHOD

It is important to discuss the limitations of the investigation method. Evaluate the method and identify any issues that could have affected the validity, accuracy, precision or reliability of the data. Sources of errors and uncertainty must also be stated in the discussion.

Once any limitations or problems in the methodology have been identified, recommend improvements on how the investigation could be conducted if repeated; for example, suggest how bias could be minimised or eliminated.

Bias

Bias may occur in any part of the investigation method, including sampling and measurements.

Bias is a form of systematic error resulting from the researcher's personal preferences or motivations. There are many types of bias, including:

- poor definitions of both concepts and variables (for example, classifying cricket pitch surfaces as slow or fast without defining 'slow' and 'fast')
- incorrect assumptions (for example, that footwear type, model and manufacturer do not affect ground reaction forces, and as a result failing to control this variable during an investigation on slip risk on different indoor and outdoor surfaces)
- errors in the investigation design and methodology (for example, taking a sample of a particular group of athletes that includes one gender more than the other in the group).

Some biases cannot be eliminated, but should at least be addressed in the discussion.

Accuracy and precision

In the discussion, evaluate the degree of accuracy and precision of the measurements for each variable of the hypothesis. Comment on the uncertainties obtained.

When relevant, compare the chosen method with any other methods that might have been selected, evaluating the advantages and disadvantages of the selected method and the effect on the results.

Reliability

When discussing the results, indicate the range of the data obtained from replicates. Explain how the sample size was selected. Larger samples are usually more reliable, but time and resources might have been scarce. Discuss whether the results of the investigation have been limited by the sample size.

The control group is important to the reliability of the investigation. A control group helps determine if a variable that should have been controlled has been overlooked and may explain any unexpected results.

Error

Discuss any source of systematic or random error and suggest ways of improving the investigation.

DISCUSSING RELEVANT PHYSICS CONCEPTS

To make the investigation more meaningful, it should be explained within the right context; i.e. using related physics ideas, concepts, theories, and models. Within this context, explain the basis for the hypothesis.

For example, if studying the impact of temperature on linear strain of a material (e.g. a rubber band), some of the contextual information to include in the discussion could be:

- the definition of linear strain
- the functions of linear strain
- the relationship between linear strain and temperature

FIGURE 11.3.3 Honest evaluation and reflection play important roles in analysing methodology.

- definitions of material behaviour (such as plastic and elastic)
- factors known to affect linear strain
- existing knowledge on the role of temperature on linear strain
- ranges of temperatures investigated and the reason these temperatures were chosen
- materials studied and the reasons for this choice
- methods of measuring the linear strain of a material.

Relating your findings to a physics concept

Once a context is established, you can use this as a framework in which to discuss whether the data supported or refuted the hypothesis. Ask questions such as:

- Was the hypothesis supported?
- Has the hypothesis been fully answered? (If not, give an explanation of why this is so and suggest what could be done to either improve or complement the investigation.)
- Do the results contradict the hypothesis? If so, why? (The explanation must be plausible and must be based on the results and previous evidence.)

Providing a theoretical context also enables comparison of the results with existing relevant research and knowledge. After identifying the major findings of the investigation, ask questions such as:

- How does the data fit with the literature?
- Does the data contradict the literature?
- Do the findings fill a gap in the literature?
- Do the findings lead to further questions?
- Can the findings be extended to another situation?

Be sure to discuss the broader implications of the findings. Implications are the bigger picture. Outlining them for the audience is an important part of the investigation. Ask questions such as:

- Do the findings contribute to or impact on the existing literature and knowledge of the topic?
- Are there any practical applications for the findings?

DRAWING EVIDENCE-BASED CONCLUSIONS

A conclusion is usually a paragraph that links the collected evidence to the hypothesis and provides a justified response to the research question.

Indicate whether the hypothesis was supported or refuted and the evidence on which this is based (that is, the results). Do not provide irrelevant information. Only refer to the specifics of the hypothesis and the research question and do not make generalisations.

Read the examples of conclusions for the following hypothesis and research question.

Hypothesis: An increase in temperature will cause an increase in linear deformation (change in length) before failure.

- Poor response to the hypothesis: Linear deformation has value y_1 at temperature 1 and value y_2 at temperature 2.
- Better response to the hypothesis: An increase in temperature from 1 to 2 produces an increase in linear deformation of z in the rubber band.

Research question: Does temperature affect the maximum linear deformation the material can withstand?

- Poor response to the research question: The results show that temperature does affect the maximum deformation of a material.
- Better response to the research question: Analysis of the results of the effect of an increase in temperature from 1 to 2 in the rubber band supports current knowledge on the effect that an increase in temperature has on increasing maximum linear deformation.

REFERENCES AND ACKNOWLEDGEMENTS

All the quotations, documents, publications and ideas used in the investigation need to be acknowledged in the 'references and acknowledgments' section to avoid plagiarism and to ensure authors are credited for their work. References and acknowledgements also give credibility to the study and allow the audience to locate information sources should they wish to study it further.

When referencing a book, include, in this order:

- author's surname and initials
- date of publication
- title
- publisher's name
- place of publication.

For example: Moran G. et al. (2017), *Pearson Physics 11*, Pearson Education, Melbourne, Victoria.

When referencing a website, include, in this order:

- author's surname and initials, or name of organisation, or title
- year website was written or last revised
- title of webpage
- date website was accessed
- website address.

For example: Wheeling Jesuit University/Center for Educational Technologies (2013), *NASA Physics Online Course: Forces and Motion*, accessed 16 June 2015, from http://nasaphysics.cet.edu/forces-and-motion.html.

11.3 Review

SUMMARY

- The discussion is the part of the investigation where the evaluation and explanation of the investigation methods and results takes place. It is the interpretation of what the results mean.
- In the discussion, the findings of the investigation need to be analysed and interpreted.
  - State whether a pattern, trend or relationship was observed between the independent and dependent variables. Describe what kind of pattern it was and specify under what conditions it was observed.
  - Were there discrepancies, deviations or anomalies in the data? If so, these should be acknowledged and explained.
  - Identify any limitations in the data collected. Perhaps a larger sample or further variations in the independent variable would lead to a stronger conclusion.
- It is important to discuss the limitations of the investigation method. Evaluate the method and identify any issues that could have affected the validity, accuracy, precision or reliability of the data. Sources of errors and uncertainty must also be stated in the discussion, and suggestions could be given as to how to reduce these errors.
- When discussing the results, indicate the range of the data obtained from replicates. Explain how the sample size was selected. Larger samples are usually more reliable, but time and resources are likely to have been scarce. Discuss whether the results of the investigation have been limited by the sample size.
- To make the investigation more meaningful, it should be explained within the right context, meaning the related physics ideas, concepts, theories and models. Within this context, explain the basis for the hypothesis.
- Indicate whether the hypothesis was supported or refuted and on what evidence this is based (that is, the results). Do not provide irrelevant information or make generalisations.

KEY QUESTIONS

1. What relationship between the variables is indicated by a sloping linear graph?
2. What relationship exists if one variable decreases as the other increases?
3. What relationship exists if both variables increase or both decrease at the same rate?
4. What might cause a sample size to be limited in an investigation?
5. Consider this investigation hypothesis: An increase in the current passing through a single resistor in an electric circuit will cause an increase in the voltage drop across the resistor.
   Improve this response to the hypothesis:
   When the current was 0.03 A, the voltage was 0.93 V and when the current was 0.05 A, the voltage was 1.81 V.

Chapter review

KEY TERMS

controlled variable
dependent variable
independent variable
mean
median
mode
outlier
personal protective equipment (PPE)
qualitative variable
quantitative variable
random error
raw data
reliability
significant figures
systematic error
uncertainty
validity
variable

11

1 What is a hypothesis and what form does it take?

2 Consider the hypothesis provided below. What are the dependent, independent and controlled variables?
Hypothesis: Releasing an arrow in archery at an angle greater or smaller than 45° will result in a shorter flight displacement (range).

3 What is the dependent variable in each of the following hypotheses?
- **a** If you push an object with a fixed mass (e.g. shot put) with a larger force, then the acceleration of that object will be greater.
- **b** The vertical acceleration of a falling object is constant.
- **c** A springboard diver rotates faster when in a tucked position than when in a stretched (layout) position.

4 List these types of hazard controls from the most effective to the least effective.
substitution, personal protective equipment, engineering controls, administrative controls, elimination, isolation

5 The speed of a toy car rolling down an inclined plane was measured 6 times. The measurements obtained (in cm s^{-1}) were 7.0, 6.5, 6.8, 7.2, 6.5, 6.5.
What is the uncertainty of the average of these values?

6 Which of the statistical measurements of mean, mode and median is most affected by an outlier?

7 What relationship between variables is indicated by a curved trend line?

8 If you hypothesise that impact force is directly proportional to drop height, what would you expect a graph of the data to look like?

9 What is meant by the ‘limitations’ of the investigation method?

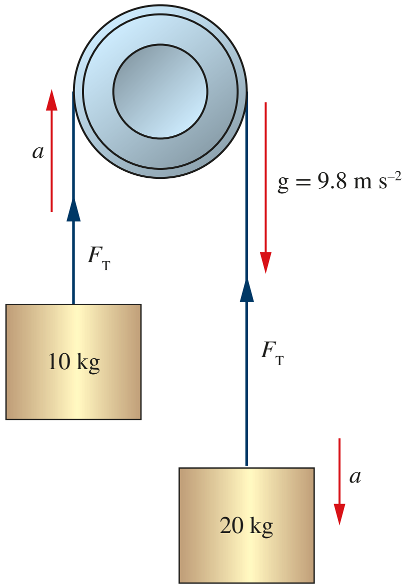
10 What is ‘bias’ in an investigation?

REVIEW QUESTIONS

Section one: Short response

1 Two students drop a lead weight from a tower and time its fall. How far does the weight travel during the second second, compared to the first second?

2 Two masses, 10 kg and 20 kg, are attached via a steel cable to a frictionless pulley, as shown in the following diagram.



a Determine the acceleration for each mass.

b What is the magnitude of the tension in the cable?

3 An engine pulls a line of rail cars along a flat track with a steady force, but instead of accelerating, the whole train travels at a constant velocity. How can this be consistent with Newton's first and second laws of motion?

4 A tow truck, pulling a car of mass 1000 kg along a straight road, causes the velocity of the car to increase from 5.00 m s^{-1} west to 10.0 m s^{-1} west in a distance of 100 m. A constant frictional force of 200 N acts on the car.

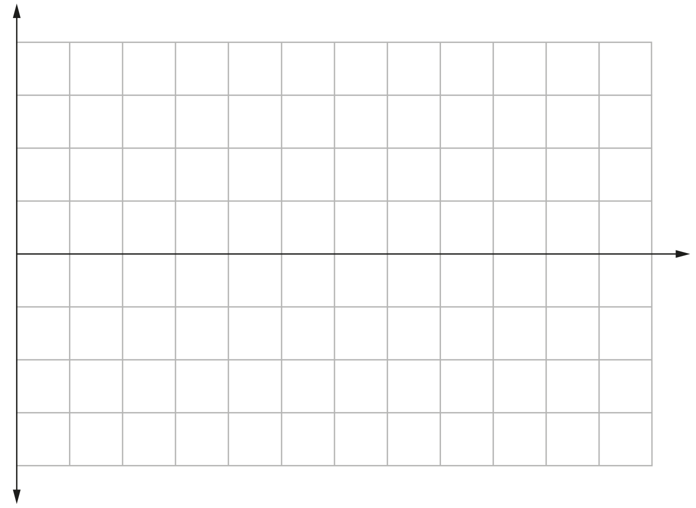
a Calculate the acceleration of the car.

b What is the resultant force acting on the car during this 100 m?

c Calculate the force exerted on the car by the tow truck.

d What force does the car exert on the tow truck?

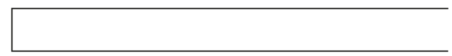
5 Copy the axes below into your workbook to complete parts (a) and (b).



a On your axes draw a labelled displacement–distance graph of a particle oscillating with a wavelength of 2.00 m and an amplitude of 3.00 cm. You must show two complete cycles.

b On another set of axes draw a labelled displacement–time graph for 1.25 cycles of a transverse wave with a period of 4.00 s and amplitude of 1.50 cm. Label this 'A'. On the same axis draw another transverse wave, also with a period of 2.00 s, but with and amplitude of 2.00 cm and out of phase from the first by 90°. Label this 'B'.

6. **a** The closed organ pipe shown below is resonating at its fifth harmonic. Draw the displacement wave occurring in the tube for this harmonic, and label the nodes with the letter 'N' and antinodes with the letter 'A'.



b If the length of the pipe in part (a) is 1.14 m, what is the fundamental frequency of vibration of this pipe?

7 Two speakers on the stage of an auditorium are 1.50 m apart and both are facing the people seated in the auditorium. The sounds from each of them are in phase and of the same amplitude. Jenny is seated 3.60 m directly in front of one of the speakers and is noticing an audible drop on certain notes from the speakers. Determine three possible frequencies that could be causing this observation.

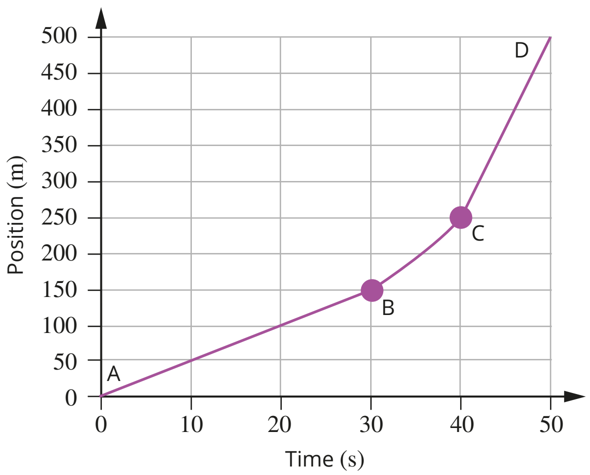
8 An earthquake produces a three-dimensional wave that flows out in all directions. It will be measured at different intensities by seismometers at various places around the world.

a Explain why the intensity measured is not the same at all of the seismometers.

b If the intensity of an earthquake wave measured by a seismometer 100 km from its source is $1.0 \times 10^6\,\mathrm{W\,m^{-2}}$, what intensity is measured 500 km from its source?

Section two: Problem solving

9 The following position versus time graph depicts the motions of a cyclist travelling east along a straight road from points A to D.



a Describe the motion of the cyclist in terms of speed.

b What was the velocity of the cyclist for the first 30 s?

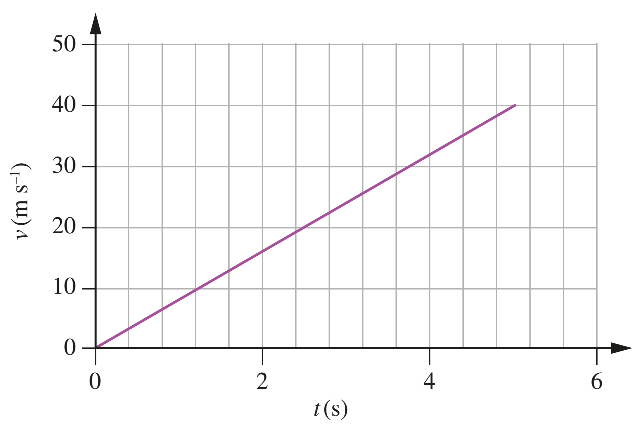
c What was the velocity of the cyclist for the final 10 s?

d Calculate the average velocity between points B and C.

e Calculate the average acceleration between points B and C.

f Calculate the average speed between points A and D.

10 The figure shows the velocity–time graph for a car of mass 2000 kg. The engine of the car is providing a constant driving force. During the 5.0 s interval the car encounters a constant frictional force of 400 N. At $t = 5.0$ s, $v = 40.0\,\mathrm{m\,s^{-1}}$.



a How much kinetic energy (in MJ) does the car have at $t = 5.0$ s?

b What is the resultant force acting on the car?

c What force is provided by the car's engine during the 5.0 s interval?

d How much work is done on the car during the 5.0 s interval?

e Determine the power output of the car's engine during the 5.0 s interval.

f How much heat energy is produced due to friction during the 5.0 s interval?

g Calculate the efficiency with which the energy provided by the engine is transformed into kinetic energy.

11 A naval cannon with mass 1.08×10^5 kg fires projectiles of mass 5.5×10^2 kg with a muzzle speed of $8.0 \times 10^2\,\mathrm{m\,s^{-1}}$. The barrel of the cannon is 20 m long, and it can be assumed that the propellant acts on the projectile for the time that it is in the barrel.

a Calculate the magnitude of the average acceleration of the projectile down the barrel.

b Using Newton's second law, calculate the average force exerted by the propellant as the projectile travels down the barrel.

c Calculate the momentum of the projectile as it leaves the barrel.

d Calculate the recoil velocity of the gun.

e Calculate the average force of the propellant from the change in momentum of the projectile.

f Calculate the average work done by the propellant on the projectile.

g Compare your answer to part (f) with the kinetic energy gained by the projectile and comment on how realistic this comparison is.

12 The didgeridoo is a musical wind instrument made and used by many Indigenous Australians in the desert regions. It is a straight hollow pipe that is often manufactured from the branch of a eucalypt tree. It varies in length but is usually around 1.5 to 2 m long. The didgeridoo has no holes, keys or valves like orchestral wind instruments; the different harmonics are simply produced by overblowing (blowing more strongly).

a Draw the pressure standing wave produced when the instrument is blown in such a way that the fifth harmonic can be heard.

b Given that the speed of sound in desert air is $336\,\mathrm{m\,s^{-1}}$, calculate the frequency of the fifth harmonic produced from this 1.50 m long didgeridoo.

c A second didgeridoo has a fundamental frequency of 106 Hz. Which is the longer of the two instruments? Justify your answer using relevant calculations.

Section three: Comprehension

13 Ionising radiation — We live in a world of radiation

Radiation is all around us all day, every day. It sustains our lives. We see because our eyes detect and analyse the radiation we call light. Infra-red radiation, whether from the sun or from a glowing fire, keeps us warm. We often cook with microwaves. Radio waves allow us to communicate sound or pictures over huge distances. Ultra-violet radiation can be used for sterilising medical equipment. All living things rely on some form of radiation for their existence.

In the twentieth century, we recognised a type of radiation known as ionising radiation. This naturally occurring radiation comes from many sources, including outer space, the sun, the rocks and soil beneath our feet, the buildings we live in, the air we breathe, the food and drink we ingest and even our own bodies. These sources combine to give us our naturally occurring background radiation dose. In Australia the average background radiation dose is approximately 1.5 mSv per year

Cosmic radiation dose rates at different altitudes

Supersonic aircraft
15,000 m
13 μSv per hour

International air travel
8,000 m
3.7 μSv per hour

La Paz, Bolivia
3900 m (highest city)
0.23 μSv per hour

Mexico City
2240 m
0.09 μSv per hour

Sea level –0.03 μSv per hour
1000 microsieverts (μSv) = 1 millisievert (mSv)
1000 millisieverts = 1 sievert (Sv)

Radiation doses

Ionising radiation and radioactive materials are widely used in medicine, industry, agriculture, environmental studies, pollution control and research. These uses benefit each of us individually and the Australian community as a whole.

Humans have increased their radiation dose through a variety of activities. One is living indoors. In surrounding ourselves with bricks and mortar, we increase the concentration of a radioactive gas called radon in the air we breathe. Radon arises naturally from the radioactive decay of uranium and thorium, normally present in rocks, soil, bricks, mortar, tiles and concrete.

Another source of radiation is medical use: – X-rays in radiography and tomography and radioactivity in nuclear medicine. Some therapeutic uses of radiation give a dose to certain organs many times higher than our annual background radiation dose.

Small extra doses of radiation occur in a number of ways. The higher you go, the less shielding the atmosphere affords from cosmic rays. On a mountain top the air may be cleaner, but the radiation dose is higher. Air travel increases radiation dose; astronauts receive even higher doses. Fallout from atmospheric nuclear testing in the 1950s and 1960s is still present in the environment. Many industries release otherwise locked-in radioactivity into the environment. This is especially true of a coal-burning plant, and to a lesser extent the fertiliser, mining and building industries.

Other common, but minor, sources of radiation are some older luminescent clocks, watches, compasses and gunsights, exit signs, certain paints and pigments, dental porcelain, re alarms, smoke detectors and television sets.

Although some radiation is capable of travelling large distances, it may be stopped by appropriate absorbers. Starlight traverses galaxies, but may be stopped by a piece of paper. Radio waves, too, are capable of travelling great distances, but may be absorbed by materials such as metals. Like light, ionising radiation travels in straight lines until absorbed or deflected. The material used to absorb ionising radiation depends on the type and energy of the radiation.

Cancer risk

Risk estimates for cancer following exposure to ionising radiation are the subject of ongoing detailed studies. Existing estimates are based largely on information of the cancer rates in the survivors of the atom bombs at Hiroshima and Nagasaki and a few other groups of people subjected to large size single or multiple doses given quickly.

Present evidence indicates that the damaging after effects of radiation exposure are greatly reduced when the dose is delivered in small amounts spread over a long time period. Nevertheless, for the purposes of radiation protection, it is assumed that any radiation dose, however small, can have some effect and protective measures are put in place for radiation workers. International studies of large groups of workers in the nuclear industry (who receive low doses spread over several years) generally agree with existing risk estimates.

Medical uses

Doses received by patients are classified separately because their exposures to radiation are justified on the grounds that such exposures pose a lesser threat to their welfare than does the risk of undiagnosed or untreated disease. No limits are set for diagnostic or therapeutic radiation exposures except that they should be as low as possible after considering risk-benet factors. In nuclear medicine, the doses from diagnostic procedures are typically in the range 1,000–10,000 μSv (average 3,300 μSv). Doses from X-ray studies are generally slightly less.

When a nuclear medicine examination is proposed for a pregnant woman, care is taken to ascertain that the examination is required for a medical condition that requires prompt therapy. For these diagnostic examinations, the risk to the mother of not performing the examination must be greater than the radiation risk to the foetus. If an examination is performed the risk to the mother and foetus is kept as low as possible.

Medical radiation sources

The average dose from medical procedures in Australia is about 800 μSv per person per year. Typical radiation doses for various medical procedures using radioisotopes are shown in for each procedure listed.

Bone Scan	4600 μSv
Thyroid Scan	2600 μSv
Lung Scan	2600 μSv
(Ventilation and perfusion)	
Liver Scan	1700 μSv
Kidney Scan	1400 μSv
Soft Tumours	40000 μSv
X-Rays (refer to table on page 5)	

All figures are based on data from Radiation Protection in Australasia, 2000.

Typical doses received during various diagnostic X-ray examinations

X-rays: (conventional)
Chest 20 µSv
Leg or foot 20 µSv
Dental 5–10 µSv
Skull 70 µSv
Barium meal 2500 µSv
Intestine 3000 µSv
Mammography 400 µSv

X-rays: (computerised tomography)

Routine head 2600 µSv

Routine abdomen 13 000 µSv

Examples of alpha, beta and gamma emitters

α emitter:

Americium-241—used in smoke detectors

β emitter:

Carbon-14—used in carbon dating

γ emitter:

Technetium 99m—diagnostic medical radio

Your average annual radiation dose

The average annual radiation dose per person per year is approximately 1500 µSv plus any exposure from medical procedures.

Some naturally occurring radiation sources are:

+50 µSv Travel and power stations—such as air travel and coal-fired power stations

+300 µSv Cosmic rays—if you live 1000 metres above sea level add 200µSv; – more if higher

+400 µSv Food and—mostly from naturally occurring radioactive potassium-40 and polonium-210. Some foods concentrate more radioactivity than others, although generally not enough to make a significant difference to this total.

+800 µSv Terrestrial radiation—long-lived radioactive materials like uranium and thorium occur in the environment. They emit ionising radiation that contributes 600 µSv a year to your average terrestrial radiation dose. This radiation comes from rocks and soils, and from building materials like bricks, mortar, concrete and tiles. radon and thoron are naturally occurring radioactive gases. Both these gases are present in the air you breathe. The major part of your average terrestrial radiation dose (200 µSv a year) therefore derives from the decay of radon and thoron in your lungs. In the open, these gases are diluted by the wind mixing them in the atmosphere. Indoors they may concentrate in still air.

Deduct 10 per cent if you live in a wooden house.

Deduct 20 per cent if you live in a tent.

Deduct 50 per cent or more if you live in the open.

Add 10 per cent or more if you live in a granite building.

Add 100 per cent or more if you keep doors and windows shut.

Add 100 per cent or more if you use bore water, especially in a hot shower. Because bore water has been underground, it contains radon that is released when the water emerges from the bore. The release is enhanced when the water is heated or divided into droplets; it is therefore most marked in a hot shower.

Source: ANSTO – Australian Nuclear Science and Technology Organisation

a What is background radiation and where does it come from?

b How do coal-fired power stations add to the background radiation?

c Why would astronauts on the International Space Station be exposed to higher doses of radiation than passengers in a commercial jet aircraft?

d Radiation can be man-made or naturally occurring in the universe. All forms of radiation can then be divided into two categories: ionising and non-ionising. Compare these two categories by completing the table below.

Type of radiation	Description	Name three types	Name one source for each
ionising			
non-ionising			

e Radon is a colourless, odourless and tasteless inert gas and is therefore chemically inactive and easily inhaled. It occurs naturally in the decay chain of uranium through thorium to lead. Uranium and thorium are among the most commonly found radioactive elements on Earth, with both having extremely long half-lives, so radon will be around for a long time yet. Radon is formed from the decay of uranium-238 in a six step process to radon-222.

- **i** The first three steps are from uranium-238 to uranium-234 via alpha, and two beta decays. Write the three equations for the decay of uranium-238 to uranium-234.
- **ii** The next three steps in the decay process all result in alpha particle emission as uranium-234 decays to radon-222. Write a single equation to represent the decay of uranium-234 to radon-222.

f Given what you know about radon, would it be safer for miners if uranium was mined underground or open cut? Explain.

g The article states, 'using bore water … or thermal springs increases your radiation dose'. Explain why this would be the case.

h A Qantas A380 flight one-way from Sydney to New York takes 15 hours. An 84.0 kg person makes 5 return flights each year. What is their absorbed dose as a result of making these flights?

i All living things undergo a carbon cycle as they recycle carbon in various compound forms. When they die, this stops and the amount of carbon at that time remains constant. The age of fossilised material can be determined using carbon dating.

- **i** Write an equation for the decay of carbon-14.
- **ii** Write an equation for the transition that occurs in the carbon-14 nucleus to produce nitrogen-14.
- **iii** Explain what carbon dating is and how this is a successful means of determining the age of fossils?

APPENDIX A SI units

PHYSICSFILE

Metric system

The metric system was originally developed in France and is known as the *Système Internationale* (SI). It was adopted in France in 1840 as the official system of units, although it had been developing in that country since 1545. It has remained in use ever since and has gradually been adopted by most other countries. It has been modified a little over the years and now, in Australia, we use SI units that have been standardised by the International Standards Organisation (IS0) since the 1960s. Some countries such as France, Italy and Spain use an earlier form of the metric system that is slightly different. The USA still measures almost everything in the old imperial units such as pounds for mass and feet for distance but, even there, scientists use the SI system of units. There are two major advantages of using the metric system. It is easier to use than other systems in that derived units are straightforward and various sizes of units are created using multiples of ten. The other very big advantage is the international nature of the standards and units. All units are standardised, making comparisons straightforward.

THE STANDARD UNITS OF MEASUREMENT

The accurate and easy measurement of quantities is essential in both everyday life and for scientific investigation. Over the centuries, many different systems of measuring physical quantities have been developed. For example, length can be measured in chains, fathoms, furlongs, yards, feet, rods and microns. Some units were based on parts of the body. The cubit was defined as the distance from the elbow to the fingertip, and so the amount of cloth that you obtained from a tailor depended on the physical size of the person selling it to you.

The metric system was established by the French Academy of Science at the time of the French Revolution (1789–1815) and is now used in most countries. This system includes units such as the metre, litre and kilogram. Countries of the British Empire adopted the British Imperial system of the mile, gallon and pound. These two systems developed independently, and their dual existence created problems in areas such as trade and scientific research. In 1960, an international committee set standard units for fundamental physical quantities. This system was an adaptation of the metric system and is known as the *Système Internationale d'Unités* (International System of Units) or SI system of units.

TABLE A.1 The SI units identify the seven fundamental quantities whose basic value is defined to a high degree of accuracy.

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

Mass

The kilogram was originally defined as the mass of 1 L of water at 4°C. This is still approximately correct, but a far more precise definition is now used. Since 1897 the measurement standard for the kilogram has been a cylindrical block of platinum–iridium alloy kept at the International Bureau of Weights and Measures in France. Australia has a copy of this standard mass at the CSIRO Division of Applied Physics in Sydney. At times it is returned to France to ensure that the mass remains accurate.

Length

The metre was originally defined in 1792 as one ten-millionth of the distance from the equator to the North Pole (approximately 10 000 km). This definition has changed a number of times since. In 1983, to give a more accurate value, the metre was redefined as the distance that light in a vacuum travels in $\frac{1}{299\,792\,458}$ second. This standard can be reproduced all over the world, as light travels at a constant speed in a vacuum.

Time

Up to 1967, time had always been based on the apparent motion of the heavens. The second was once defined in terms of the motion of the Sun. Until 1960, one second was defined as $\frac{1}{60}$ of $\frac{1}{60}$ of $\frac{1}{24}$ of an average day in 1900. This reflected the rate of the Earth's rotation on its axis; however, its rotation is not quite uniform. In 1967, a more accurate definition was adopted—one not based on the motion of the Earth. One second is now defined as the time required for a caesium-133 atom to undergo 9 162 631 770 vibrations. These vibrations are stimulated by an electric current and are extremely stable, allowing this standard to be reproduced all over the world.

DERIVED UNITS

As well as the seven fundamental quantities, a wide variety of other physical quantities can be measured. You may have encountered some of these, such as frequency, velocity, energy and density, already. A derived quantity is defined in terms of the fundamental quantities. For example, the SI unit for area is square metres (m^2).

TABLE A.2 Some derived SI quantities and their units.

Quantity	SI unit	SI unit symbol	Equivalent unit
velocity	metres per second	$m\,s^{-1}$	—
acceleration	metres per second per second	$m\,s^{-2}$	—
frequency	hertz	Hz	s^{-1}
force	newton	N	$kg\,m\,s^{-2}$
energy/work	joule	J	$kg\,m^2\,s^{-2}$

Understanding measurement

MEASUREMENT AND UNITS

In every area of physics we have attempted to quantify the phenomena we study. In practical demonstrations and investigations we generally make measurements and process those measurements in order to come to some conclusions. Scientists have a number of conventional ways of interpreting and analysing data from their investigations. There are also conventional ways of writing numerical measurements and their units.

Correct use of unit symbols

The correct use of unit symbols removes ambiguity, as symbols are recognised internationally. The symbols for units are not abbreviations and should not be followed by a full stop unless they are at the end of a sentence.

Upper-case letters are not used for the names of any physical quantities of units. For example, we write newton for the unit of force, while we write Newton if referring to someone with that name. Upper-case letters are only used for the *symbols* of the units that are named after people. For example, the unit of energy is joule and the symbol is J. The joule was named after James Joule who was famous for studies into energy conversions. The exception to this rule is 'L' for litre. We do this because a lower-case 'l' looks like the numeral '1'. The unit of distance is metre and the symbol is m. The metre is not named after a person.

The product of a number of units is shown by separating the symbol for each unit with a dot or a space. Most teachers prefer a space but a dot is perfectly correct. The division or ratio of two or more units can be shown in fraction form, using a slash, or using negative indices. Most teachers prefer negative indices. Prefixes should not be separated by a space.

TABLE B.1 Some examples of the use of symbols for derived units.

Preferred	Correct also	Wrong
m s^{-2}	m.s^{-2} m/s^2	ms^{-2}
kW h	kW.h	kWh k Wh
kg m^{-3}	kg.m^{-3} kg/m^3	kgm^{-3}
µm		µ m
N m	N.m	Nm

Units named after people can take the plural form by adding an 's' when used with numbers greater than one. Never do this with the unit symbols. It is acceptable to say 'two newtons' but wrong to write 2 Ns. It is also acceptable to say 'two newton'.

Numbers and symbols should not be mixed with words for units and numbers. For example, twenty metres and 20 m are correct while 20 metres and twenty m are incorrect.

Scientific notation

To overcome confusion or ambiguity, measurements are often written in scientific notation. Quantities are written as a number between one and ten and then multiplied by an appropriate power of ten. Note that 'scientific notation', 'standard notation' and 'standard form' all have the same meaning.

Examples of some measurements written in scientific notation are:

$0.054\,m = 5.4 \times 10^{-2}\,m$

$245.7\,J = 2.457 \times 10^{2}\,J$

$2080\,N = 2.080 \times 10^{3}\,N$ or $2.08 \times 10^{3}\,N$

You should be routinely using scientific notation to express numbers. This also involves learning to use your calculator intelligently. Scientific and graphics calculators can be put into a mode whereby all numbers are displayed in scientific notation. It is useful when doing calculations to use this mode rather than frequently attempting to convert to scientific notation by counting digits on the calculator display. It is quite acceptable to write all numbers in scientific notation, although most people prefer not to use scientific notation when writing numbers between 0.1 and 1000.

An important reason for using scientific notation is that it removes ambiguity about the precision of some measurements. For example, a measurement recorded as 240 m could be a measurement to the nearest metre; that is, somewhere between 239.5 m and 240.5 m. It could also be a measurement to the nearest ten metres; that is, somewhere between 235 m and 245 m. Writing the measurement as 240 m does not indicate either case. If the measurement was taken to the nearest metre, it would be written in scientific notation as $2.40 \times 10^{2}\,m$. If it was taken to the nearest ten metres only, it would be written as $2.4 \times 10^{2}\,m$.

FIGURE B.1 A scientific calculator.

PREFIXES AND CONVERSION FACTORS

Conversion factors should be used carefully. You should be familiar with the prefixes and conversion factors in Table B.2. The most common mistake made with conversion factors is multiplying rather than dividing. Some simple strategies can save you this problem. Note that the table gives all conversions as a multiplying factor.

TABLE B.2 Prefixes and conversion factors.

Multiplying factor		Prefix	Symbol
1 000 000 000 000	10^{12}	tera	T
1 000 000 000	10^{9}	giga	G
1 000 000	10^{6}	mega	M
1 000	10^{3}	kilo	k
0.01	10^{-2}	centi	c
0.001	10^{-3}	milli	m
0.000 001	10^{-6}	micro	μ
0.000 000 001	10^{-9}	nano	n
0.000 000 000 001	10^{-12}	pico	p

Do not put spaces between prefixes and unit symbols. It is important to give the symbol the correct case (upper or lower case). There is a big difference between 1 mm and 1 Mm.

There is no space between prefixes and unit symbols. For example, one-thousandth of an ampere is given the symbol mA. Writing it as m A is incorrect. The space would mean that the symbol is for a derived unit—a metre ampere.

Worked example B1

The diameter of a cylindrical piece of copper rod was measured at 24.8 mm with a vernier caliper. Its length was measured at 35 cm with a tape measure.
a Find the area of cross-section in m^2.
b Find the volume of the copper rod in m^3.
Answer
a The area of cross-section is πr^2. The radius is calculated by dividing the diameter by two. Hence the radius is 12.4 mm. To calculate the area in m^2, first halve the diameter and convert it to metres. The radius is $\frac{24.8}{2} = 12.4\,\text{mm} = 12.4 \times 10^{-3}\,\text{m}$. The radius is not written in scientific notation. This is not necessary. All you need to do is multiply by the appropriate factor. The conversion factor for mm to m is 10^{-3}. Just multiply by the conversion factor and don't bother to rewrite the result in scientific notation. This is because it is only going to be used in a calculation and is not a final result. The area of cross-section is $\pi r^2 = \pi(12.4 \times 10^{-3})^2 = 4.8 \times 10^{-4}\,\text{m}^2$.
b The volume is $\pi r^2 h$, where h is the length of the cylinder. The length is $35\,\text{cm} = 35 \times 10^{-2}\,\text{m}$. Hence the volume is $\pi(12.4 \times 10^{-3})^2(35 \times 10^{-2}) = 1.7 \times 10^{-4}\,\text{m}^2$.

Worked example B2

a A car is traveling at 110 km h^{-1}. How fast is this in m s^{-1}?
b Convert 35 miles per hour to metres per second. A mile is approximately 1600 m.
Answer
a 110 km h^{-1} is 110×10^3 metres per 3600 s. $\frac{110 \times 10^3}{3600} = 30.6$ Hence $110\,\text{km h}^{-1} = 30.6\,\text{m s}^{-1}$.
b 35 miles per hour is 35×1600 metres per 3600 s. $\frac{35 \times 1600}{3600} = 15.6$ Hence 35 mph = $16\,\text{m s}^{-1}$.

DATA

Physicists and physics students collect, analyse and interpret experimental data. In fact, you will do this when you conduct your Practical Investigation. Working with data requires a good understanding of the meaning and limitations of measurement.

Accuracy and precision

Two very important aspects of any measurement are accuracy and precision. Accuracy and precision are not the same thing. The distinction between the two ideas is only hard to grasp because the two words are defined in a similar way in the dictionary. We often hear the words used together, and in general conversation they tend to be used interchangeably.

Instruments are said to be *accurate* if they truly reflect the quantity being measured. For example, if a tape measure is correctly manufactured it can be used to measure lengths accurately to the nearest centimetre.

Imagine that the tape measure is accidentally stretched during the manufacturing process, as shown in Figure B.2. It would still be used to measure length to the nearest centimetre but all measurements would be wrong. It would be inaccurate.

Suppose an accurate ruler had 3 cm snapped off the end, as shown in Figure B.3. It would now give readings all too large by 3 cm if no allowance were made for the missing piece. This ruler measure would be inaccurate.

In these two examples, the tape measure or ruler is used to measure to the nearest centimetre but is inaccurate. Inaccurate means just plain wrong. Instruments are said to be *precise* if they can differentiate between slightly different quantities. Precision refers to the fineness of the scale being used.

Consider the metre rule, the tape measure and the measuring wheel used to mark out sports fields. All three measure distance. All three can be accurate. The metre rule is more *precise* because it measures to the nearest millimetre, the tape measure has less precision due to measuring only to the nearest centimetre, while the wheel measures only to the nearest metre (Figure B.4).

The tape measure is a more precise instrument than the measuring wheel. Suppose two distances of 2673 mm and 2691 mm are being measured with these two instruments. Each distance would be measured as 3 m, to the nearest metre, by the wheel. They would be measured differently as 2.67 m and 2.69 m, to the nearest centimetre, by the tape measure. The tape measure is more precise because it has a finer scale. You might also say that it has greater resolution. The measuring wheel has such low precision that it can't be used to measure which of the two distances is greater or smaller. Measuring instruments with less precision give measurements that are less certain. The uncertainty in the measurement is due to a coarser scale. The measuring wheel gives less certain measurements than the tape measure even though both instruments may be equally accurate.

All measurements have some amount of *uncertainty*, due to the precision of the instrument that does the measuring. (Note that in Chapter 10 the uncertainty due to collecting a range of data was analysed. This section deals with uncertainty due to precision.) The uncertainty is generally one half of the finest scale division on the measuring instrument. The measuring wheel has an uncertainty of 0.5 m. The metre rule has an uncertainty of 0.5 mm. The tape measure has an uncertainty of 0.5 cm. An electronic balance set to measure grams to two decimal places has an uncertainty of 0.005 g.

Sometimes this uncertainty is referred to as error. It is not error, in that it is not a mistake or something wrong. All measuring instruments have limited precision and, in general, the uncertainty is half of the smallest scale division on the instrument.

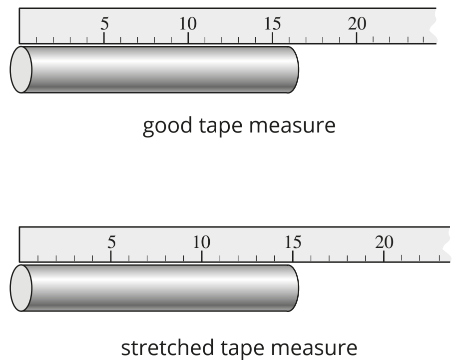


FIGURE B.2 The diagram shows that a correctly manufactured tape measure correctly measures the cylinder to be 16 cm long while the stretched tape measure gives a wrong measurement of 15 cm. The stretched tape measure is inaccurate.

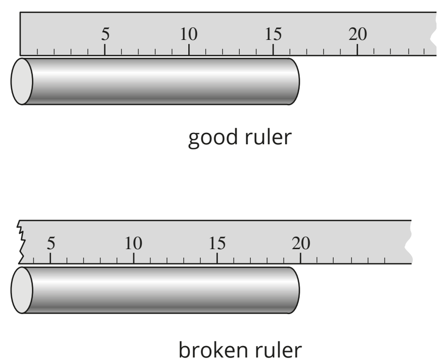


FIGURE B.3 The diagram shows that an undamaged ruler correctly measures the cylinder to be 16 cm long while the broken ruler gives a wrong measurement of 19 cm. The broken ruler is inaccurate but equally as precise as the unbroken ruler.

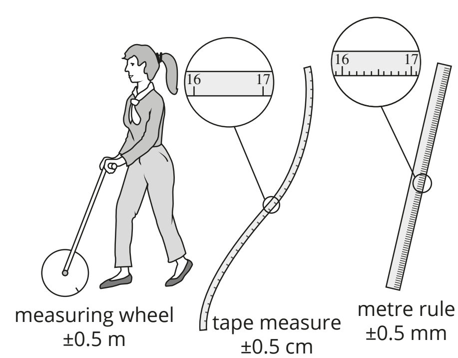


FIGURE B.4 The measuring wheel has low precision and only measures to the nearest metre. It has an uncertainty of 0.5 m. The tape measure has more precision and has an uncertainty of 0.5 cm or 0.005 m. The metre rule has an uncertainty of 0.5 mm or 0.0005 m.

The uncertainty is the measure of the precision of an instrument. It is not related to accuracy. A micrometer screw gauge, which measures length to the nearest one-hundredth of a millimetre and hence is very precise, may not be accurate. Usually they are, but if one has been badly manufactured or bent by being over-tightened repeatedly it most likely will be inaccurate. But its precision will still be ±0.000 005 m, or half of one-hundredth of a millimetre.

The uncertainty gives the range in which a measurement falls. If you measured the length of a stick with a metre rule then you would get a measurement 'plus or minus' half a millimetre.

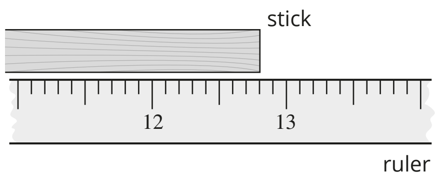


FIGURE B.5 A stick anywhere between 127.5 mm and 128.5 mm would be recorded as having a length of 128 mm if measured by a metre rule with a scale division of 1 mm. Conversely, a measurement recorded as 128 mm could be of an object of length anywhere between 127.5 mm and 128.5 mm.

Any stick between 127.5 mm and 128.5 mm long would be measured as 128 mm to the nearest millimetre (refer to Figure B.5). We would record this as 128 ± 0.5 mm.

When using an analogue scale, you might think that you can 'judge by eye' fractions of a scale division and hence get greater precision than half a scale division. You should be able to judge to the nearest half a scale division. You might think you can judge to the nearest tenth of a division. You can't. Research shows that despite the fact that people try to judge the spaces between scale divisions to better than half a division, as soon as this is done, inconsistent measurements are obtained. That is, different people get different measurements of the same thing.

The best judgement you can definitely claim is one half of a scale division. The uncertainty we will still assume, however, is a full half-scale division. Hence, you might measure another stick, one that has a length somewhere between 154 mm and 155 mm, as 154.5 ± 0.5 mm.

Of course, you don't have the option of adding an extra decimal place containing a 0 or a 5 if you are using a digital instrument.

The uncertainty can be recorded as the *absolute* uncertainty as we have done above. The absolute uncertainty is the actual uncertainty in the measurement. In this case it is 0.5 mm. It is often useful to write the uncertainty as a *percentage*: 0.5 mm is 0.32% of 154.5. Hence, the above length would be recorded as 154.5 mm ± 0.32%.

Percentage uncertainty is also called *relative* uncertainty. It is the size of the uncertainty relative to the size of the measured quantity.

PHYSICSFILE

Many people use the term 'error' to refer to uncertainty and many other things. The problem with referring to uncertainty as error is that it is not actually error. Things that are a normal consequence of the limitations of measuring instruments must happen, and are not mistakes. If they are not mistakes or 'something gone wrong' then it makes no sense to call them errors.

Errors are the factors that limit the accuracy of your results. For example, if you perform a calorimetry experiment and do not use a good enough insulator, you will get inaccurate results due to heat losses to the environment. This will contribute to the error in your measurement. Suppose you measured the refraction of light in glass but did not place the protractor in the correct place when measuring angles. This would also cause error.

Many different things can contribute to experimental error. Some are unavoidable. Some are factors in the design of the experiment. Good experimental design seeks to eliminate or at least minimise potential sources of error.

Never quote 'human error' as a source of error. Your data should be examined carefully and mistakes eliminated or at least ignored. So-called human errors, or lack of care, have no place in your experimental work. If you make mistakes then you should repeat the measurements.

Estimating the uncertainty in a result

An experiment or a measurement exercise is not complete until the uncertainties have been analysed. Chapter 11 explained how uncertainties were treated when a range of data had been collected and then averaged out. It is also important to explain how uncertainty due to the precision of instruments affects results.

The following three processes are used for estimating uncertainty in calculations due to the precision of instruments. They are demonstrated in Worked example B3.

- When adding or subtracting data, add the absolute uncertainties.
- When multiplying or dividing data, add the percentage uncertainties.
- When raising data to power *n*, multiply the percentage uncertainty by *n*.

In Worked example B3, the analysis of uncertainty reveals the *precision* of an experimental result.

Worked example B3

You might have measured the specific heat capacity of a metal. You could have calculated your result using:

$$c_{metal} = \frac{c_{water} m_{water} \Delta T_{water}}{m_{metal} \Delta T_{metal}}$$

Suppose you had the following data included in your table.

Quantity		Absolute uncertainty	% uncertainty
c_{water}	$4180\,\mathrm{J\,kg^{-1}\,K^{-1}}$	$5\,\mathrm{J\,kg^{-1}\,K^{-1}}$	0.120
m_{water}	$72.5 \times 10^{-3}\,\mathrm{kg}$	$0.05 \times 10^{-3}\,\mathrm{kg}$	0.069
ΔT_{water}	5°C	1°C*	20
m_{metal}	$87.3 \times 10^{-3}\,\mathrm{kg}$	$0.05 \times 10^{-3}\,\mathrm{kg}$	0.057
ΔT_{metal}	72°C	1°C*	1.389

*Note that the ΔT values have an absolute uncertainty of 1°C because they are calculated by subtracting one temperature measurement from another.

You would calculate as follows:

$$\begin{aligned} c_{metal} &= 241\,\mathrm{J\,kg^{-1}\,K^{-1}} \\ \text{Uncertainty (\%)} &= 0.120 + 0.069 + 20 + 0.057 + 1.389 \\ &= 21.6\% \end{aligned}$$

Hence, you would obtain the following result:

$$\begin{aligned} c_{metal} &= 241\,\mathrm{J\,kg^{-1}\,K^{-1}} \pm 21.6\% \\ c_{metal} &= 241 \pm 52\,\mathrm{J\,kg^{-1}\,K^{-1}} \end{aligned}$$

Once you have done all of this you can consider the relative success of your measurement exercise.

Your result is:

$$189\,\mathrm{J\,kg^{-1}\,K^{-1}} \le c_{metal} \le 293\,\mathrm{J\,kg^{-1}\,K^{-1}}$$

If measurements by other people, such as the constants published in data books, fall within this range then you can conclude that your experiment is consistent with established values. That is, within the precision of your technique, there are probably no significant errors although the final measurement is rather imprecise in this case. We might say that it is accurate within the limitations of the equipment.

PHYSICSFILE

In some classes, students are instructed to quote all results to two decimal places or to three significant figures. You should be able to see from Worked example B3 that these rules are not absolutely correct when applied to real data. For ordinary calculations in assignments, tests and examinations, you might just give your answers to three figures.

If a calculation is done in several stages then you should not round off any intermediate results. This will add rounding error to your calculations. Use the memory on your calculator so that there is no rounding until the end of your calculation.

You are also now in a position to refine the experiment by reducing the larger uncertainties. In this case, the largest uncertainty was in the temperature change for the water. Hence, it would not be very helpful to measure the masses to greater precision because the limit to precision in this activity would be the temperature differences. Getting greater precision in the temperature changes would be a useful refinement.

You could consider ways of getting larger temperature changes in the water and hence obtain a smaller percentage uncertainty in the temperature change. Alternatively, you might consider ways of measuring the temperatures to greater precision.

If your measurement range does not include the result you expect, you should think about the origin of the errors. In other words, if you are sure that c_{metal} is less than 189 J kg^{-1} K^{-1} or more than 293 J kg^{-1} K^{-1} then there must be some error in your experimental technique or more uncertainty than you realised.

When reviewing an experiment or a measurement exercise, it is a good idea to consider both errors *and* uncertainties.

Significant figures

The number of significant figures in a measurement is simply the number of digits used when the number is written in scientific notation. (Note: Significant figures were explained in Chapter 11.) Your calculator usually has eight or ten digits in the display of the answer for a calculation, but most of them are meaningless. You must round off your answer appropriately.

Consider the result of the experiment described in Worked example B3. It would make no sense to quote the result to two decimal places (or five significant figures) when clearly the precision of the experiment gives less than three significant figures.

Calculated results never have more significant figures than the original data and might have fewer than the original data. If you are not doing a full analysis of the uncertainties, it is customary to give your answers to the same number of significant figures as the least precise piece of data. For example, in Worked example B3, the least precise data is the change in temperature of the water, which has only a single digit. The value for the specific heat might then be quoted simply as 2×10^2 J kg^{-1} K^{-1}, but doing the full calculation of the uncertainty in the result is much more informative.

GRAPHICAL ANALYSIS OF DATA

A major problem with doing a calculation from just one set of measurements is that a single incorrect measurement can significantly affect the result. Scientists like to take a large amount of data and observe the trends in that data. This gives more precise measurements and allows scientists to recognise and eliminate problematic data.

Physicists commonly use graphical techniques to analyse a set of data. In this section, the basic techniques that they use will be outlined and a general method for using a set of data that fits a known mathematical relationship will be developed.

Linear relationships

Some relationships studied in physics are linear, that is a straight line, while others are not. It is possible to manipulate non-linear data so that a linear graph reveals a measurement. Linear relationships and their graphs are fully specified with just two numbers: gradient, m, and vertical axis intercept, c. In general, linear relationships are written:

$$y = mx + c$$

The gradient, m, can be calculated from the coordinates of two points on the line:

$$m = \frac{\text{rise}}{\text{run}}$$

$$= \frac{y_2 - y_1}{x_2 - x_1}$$

where (x_1, y_1) and (x_2, y_2) are any two points on the line. Don't forget that m and c have units. Omitting these is a common error.

PHYSICSFILE

Graphs

When analysing data from a linear relationship, it is first necessary to obtain a graph of the data and an equation for the line that best fits the data. This line of best fit is often called the regression line. The entire process can be done on paper but most people will use a computer spreadsheet, the capabilities of a scientific or a graphics calculator, or some other computer-based process. In what follows, it is not assumed that you are using any particular technology.

If you are plotting your graph manually on paper then proceed as follows:

1. Plot each data point on clearly labelled, unbroken axes.
2. Identify and label but otherwise ignore any suspect data points.
3. Draw, by eye, the 'line of best fit' for the points. The points should be evenly scattered either side of the line.
4. Locate the vertical axis intercept and record its value as 'c'.
5. Choose two points on the line of best fit to calculate the gradient. Do not use two of the original data points as this will not give you the gradient of the line of best fit.
6. Write $y = mx + c$, replacing x and y with appropriate symbols, and use this equation for any further analysis.

If you are using a computer or a graphics calculator then proceed as follows:

1. Plot each data point on clearly labelled, unbroken axes.
2. Identify suspect data points and create another data table without the suspect data.
3. Plot a new graph without the suspect data. Keep both graphs, as you don't actually discard the suspect data, but do eliminate it from the analysis.
4. Plot the line of best fit—the regression line. The manner in which you do this depends on the model of calculator or the software being used.
5. Compute the equation of the line of best fit that will give you values for m and c.
6. Write $y = mx + c$, replacing x and y with appropriate symbols, and use this equation for any further analysis.

Worked example B4

Some students used a computer with an ultrasonic detector to obtain the speed–time data for a falling tennis ball. They wished to measure the acceleration of the ball as it fell. They assumed that the acceleration was nearly constant and that the relevant relationship was $v = u + at$, where v is the speed of the ball at any given time, u was the speed when the measurements began, a is the acceleration of the ball and t is the time since the measurement began.

Their computer returned the following data:

Time (s)	Speed (m s^{-1})
0.0	1.25
0.1	2.30
0.2	3.15
0.3	4.10
0.4	5.25
0.5	6.10
0.6	6.95

Find their experimental value for acceleration.

Solution

The data is assumed linear, with the relationship $v = u + at$, which can be thought of as being $v = at + u$, which makes it clear that putting v on the vertical axis and t on the horizontal axis gives a linear graph with gradient a and vertical intercept u. A graph of the data is shown in Figure B.6.

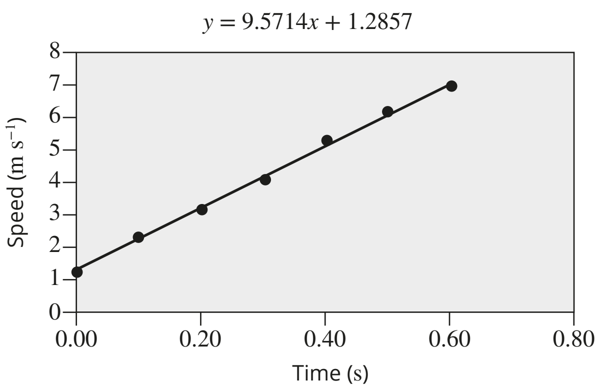


FIGURE B.6 Speed–time profile for a falling tennis ball.

This graph of the data was created on a computer spreadsheet. The line of best fit was created mathematically and plotted. The computer calculated the equation of the line. Graphics calculators can also do this.

A scientific calculator, graphics calculator or spreadsheet gives the regression line as $y = 9.5714x + 1.2857$. If this is rearranged and the constants are suitably rounded, the equation is $v = 1.3 + 9.6t$. This indicates that the ball was moving at 1.3 m s^{-1} at the commencement of data collection and accelerating at 9.6 m s^{-2}.

Manipulating non-linear data

Suppose you were examining the relationship between two quantities B and d and had good reason to believe that the relationship between them is

$$B = \frac{k}{d}$$

where k is some constant value. Clearly, this relationship is non-linear and a graph of B against d will not be a straight line. By thinking about the relationship it can be seen that in 'linear form':

$$\begin{array}{ccccc} B & = & k & \frac{1}{d} & \\ \uparrow & & \uparrow & \uparrow & \\ y & = & m & x & + c \end{array}$$

A graph of B (on the vertical axis) against $\frac{1}{d}$ (on the horizontal axis) will be linear. The gradient of the line will be k and the vertical intercept, c, will be zero. The line of best fit would be expected to go through the origin because, in this case, there is no constant added and so c is zero.

In the above example, a graph of the raw data would just show that B is larger as d is smaller. It would be impossible to determine the mathematical relationship just by looking at a graph of the raw data.

A graph of raw data will not give the mathematical relationship between the variables, but it can give some clues. The shape of the graph of raw data may suggest a possible relationship. Several relationships may be tried and then the best is chosen. Once this is done, it is not proof of the relationship but, possibly, strong evidence.

When an experiment involves a non-linear relationship, the following procedure is followed:

1 Plot a graph of the original raw data.

2 Choose a possible relationship based on the shape of the initial graph and your knowledge of various mathematical and graphical forms.

3 Work out how the data must be manipulated to give a linear graph.

4 Create a new data table.

Then follow the steps given in the Physics file on page 10. It may be necessary to try several mathematical forms to find one that seems to fit the data.

Worked example B5

Some students were investigating the relationship between current and resistance for a new solid-state electronic device. They obtained the data shown in the table.

According to the theory they had researched, the students believed that the relationship between I and R is $R = dI^3 + g$, where d and g are constants.

By appropriate manipulation and graphical techniques, find their experimental values for d and g. The following steps should be used:

a Plot a graph of the raw data.

b Work out what you would have to graph to get a straight line.

c Make a new table of the manipulated data.

d Plot the graph of manipulated data.

e Find the equation relating I and R.

Current, I (A)	Resistance, R (Ω)
1.5	22
1.7	39
2.2	46
2.6	70
3.1	110
3.4	145
3.9	212
4.2	236

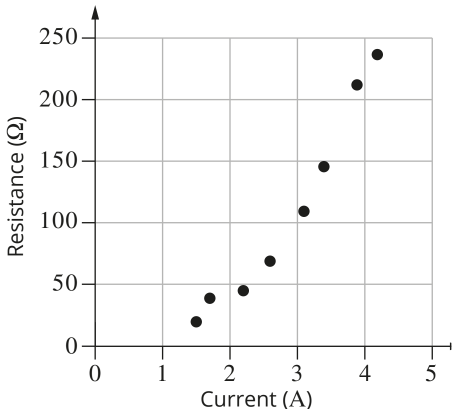


FIGURE B.7 Current–resistance graph of device.

Solution

a Figure B.7 shows the graph obtained using a spreadsheet.

It might be argued that the second piece of data is suspect. The rest of this solution supposes the students chose to ignore this piece of data.

b You can see what to graph if you think of the equation like this:

$$\begin{array}{ccccccc} R & = & d & I^3 & + & g \\ \uparrow & & \uparrow & \uparrow & & \uparrow \\ y & = & m & x & + & c \end{array}$$

A graph of R on the vertical axis and I^3 on the horizontal axis would have a gradient equal to d and a vertical axis intercept equal to g.

c The data is manipulated by finding the cube of each of the values for current.

Current cubed, I^3 (A^3)	Resistance, R (Ω)
3.38	22
10.65	46
17.58	70
29.79	110
39.30	145
59.32	212
74.09	236

d The graph in Figure B.8 was obtained from the spreadsheet.

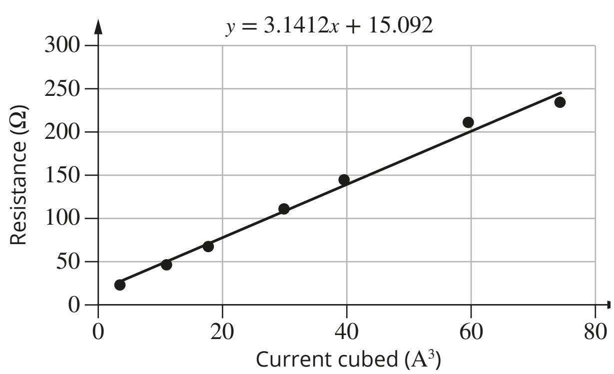


FIGURE B.8 Current–resistance characteristic (manipulated data).

e The regression line has the equation $y = 3.1x + 15.1$, so the equation relating I and R is $R = 3.1I^3 + 15.1$. Hence, the value of d is 3.1 $\Omega\ A^{-3}$ and the value of g is 15.1 Ω.

Maths Skills

1. Transforming decimal notation to scientific notation

Scientists use scientific notation to handle very large and very small numbers.

For example, instead of writing 0.000 000 035, scientists would write 3.5×10^{-8}.

A number in *scientific notation* (also called standard form or power of ten notation) is written as:

$$a \times 10^n$$

where a is a number equal to or greater than 1 and less than 10, that is $1 \leq a < 10$.

n is an integer (a positive or negative whole number).

n is the power that 10 is raised to and is called the index value.

To transform a very large or very small number into scientific notation:

Write the original number as a decimal number greater than or equal to 1 but less than 10.

Multiply the decimal number by the appropriate power of 10.

The index value is determined by counting the number of places the decimal point needs to be moved to form the original number again.

If the decimal point is moved n places to the right, n will be a positive number. For example:

$$51 = 5.1 \times 10^1$$

If the decimal point is moved n places to the left, n will be a negative number. For example:

$$0.51 = 5.1 \times 10^{-1}$$

You will notice from these examples that when large numbers are written in scientific notation, the 10 has a positive index value. When very small numbers are written in scientific notation, the 10 has a negative index value.

Practice questions

1 Match each number with its correct scientific notation.

Number	Scientific notation
0.002	2×10^3
2000	1.234×10^{-1}
0.1234	2×10^{-3}
12.34	1.234×10^1
123.4	1.234×10^2

2 Write 7.009×10^{-4} using decimal notation.

2. Identifying significant figures

When giving an answer to a calculation it is important to take note of the number of significant figures that you use.

You should give an answer that is as accurate as possible. However, an answer can't be more accurate than the data or the measuring device used to calculate it. For example, if a set of scales that measures to the nearest gram shows that an object has a mass of 56 g, then the mass should be recorded as 56 g, not 56.0 g. This is because you do not know whether it is 56.0 g, 56.1 g, 56.2 g or 55.8 g.

56 is a number with 2 significant figures. Recording to 3 significant figures (e.g. 56.0 g or 55.8 g) would not be scientifically 'honest'.

If this mass of 56 g is used to calculate another value it would also not be 'honest' to give an answer that has more than 2 significant figures.

Determining to what number of significant figures to give an answer to depends on what kind of calculation you are doing.

If you are multiplying or dividing, use the smallest number of significant figures provided in the initial values.

If you are adding or subtracting, use the smallest number of decimal places provided in the initial values.

Working out the number of significant figures

The following rules should be followed to avoid confusion in determining how many significant figures are in a number.

1 All non-zero digits are always significant. For example, 21.7 has 3 significant figures.
2 All zeroes between two non-zero digits are significant. For example, 3015 has 4 significant figures.
3 A zero to the right of a decimal point and following a non-zero digit is significant. For example, 0.5700 has 4 significant figures.
4 Any other zero is not significant, as it will be used only for locating decimal places. For example, 0.005 has just 1 significant figure.

Practice questions

1 Which of the following is written to 2 significant figures?
 A 30.1
 B 0.000 40
 C 0.5
 D 5.12

2 George multiplied 1.22 by 1.364. Which of the options below shows the result of this multiplication with the correct number of significant figures?
 A 1.66
 B 1.664
 C 1.65
 D 1.7

3 How can 41 be written to 4 significant figures?
 A 00.41
 B 4100
 C 41.00
 D 4.100

4 Alex is getting ready to go for a bike ride. Alex's mass is 65.3 kg. The bicycle has a mass of 12.92 kg.
 a Calculate the combined mass of Alex and the bicycle. Give your answer to the correct number of significant figures.
  A 78
  B 78.2
  C 78.22
  D 78.3
 b Using the combined mass calculated in part (a) above, and the formula $F_{net} = ma$, calculate the force Alex needs to apply to achieve an acceleration of 1.250 m s^{-1}. Give your answer to the correct number of significant figures.

3. Calculating percentages

Scientists use percentages to express a ratio or fraction of a quantity.

To express one quantity as a percentage of another, use the second quantity to represent 100%.

For example, expressing 6 as a percentage of 24 is like saying '6 is to 24 as x is to 100':

$$\frac{6}{24} = \frac{x}{100}$$
$$x = \frac{6}{24} \times 100$$
$$= 25\%$$

To calculate a percentage of a quantity, the percentage is expressed as a decimal then multiplied by the quantity.

For example, to calculate 40% of 20:

$$x = \frac{40}{100} \times 20$$
$$= 0.4 \times 20$$
$$= 8$$

Practice questions

1 What is 9 as a percentage of 12?
- **A** 25%
- **B** 50%
- **C** 75%
- **D** 30%

2 What is 25% of 24?
- **A** 6.6
- **B** 6
- **C** 5
- **D** 0.5

3 Which of the following values expresses 15 as a percentage of 120?
- **A** 8%
- **B** 5%
- **C** 12.5%
- **D** 0.125%

4. Converting between percentages and fractions

To write a percentage as a fraction, divide the percentage by 100.

For example:

$$25\% = \frac{25}{100}$$
$$= \frac{1}{4}$$

$\frac{25}{100}$ is not the simplest form of this fraction. If you divide both the numerator and the denominator by 25 (their highest common factor) then the fraction simplifies to $\frac{1}{4}$.

Whenever you give a fraction as an answer, always try and simplify it by dividing the numerator and denominator by the highest common factor.

To write a fraction as a percentage, multiply the fraction by 100%. In many cases it is easier to convert the fraction to a decimal number first.

For example:

$$\frac{1}{4} = 0.25 \times 100$$
$$= 25\%$$

The value of the fraction or percentage has not changed. It is just being represented in a different way.

Practice questions

1 Choose the option that expresses $\frac{1}{5}$ as a percentage.
A 25%
B 20%
C 30%
D 50%

2 Match each percentage with its corresponding fraction.

Percentage	Fraction
0.2%	$\frac{7}{20}$
2.5%	$\frac{7}{40}$
17.5%	$\frac{1}{40}$
35%	$\frac{111}{250}$
44.4%	$\frac{1}{500}$

5. Changing the subject of an equation

Scientists use equations to represent relationships between variables. In an equation like $A = \pi r^2$, A is called the subject of the equation.

Sometimes the subject of the equation has to be changed in order to express the relationship in a more useful way. For example, if you need to find the radius of a circle, you would want r to be the subject of the equation above.

To change the subject of a simple equation, transpose the equation to leave the new subject on its own. In the example above, the equation needs to read

$r = \ldots$

Keep the equation balanced by performing the same operation to both sides of the equation to cancel operations being performed on the desired subject. Inverse operations (the opposite operation; for example dividing is the inverse of multiplying) will allow cancelling.

For example, make r the subject of the equation $A = \pi r^2$.

1 Divide both sides of the equation by π.

$$A = \pi r^2$$

$$\frac{A}{\pi} = \frac{\pi r^2}{\pi}$$

The π in the numerator and denominator on the right side of the equation cancel out, giving

$$\frac{A}{\pi} = r^2$$

2 To cancel the squaring operation of r, take the square root of both sides of the equation.

$$\frac{A}{\pi} = r^2$$

$$\sqrt{\frac{A}{\pi}} = \sqrt{r^2}$$

The square and square root on the right side of the equation cancel out, giving

$$\sqrt{\frac{A}{\pi}} = r$$

3 r is now the subject of the equation.

$$r = \sqrt{\frac{A}{\pi}}$$

Practice questions

1 Rearrange the formula $A = \frac{2}{3}R$ to make R the subject.

A $R = \frac{2A}{3}$

B $R = \frac{A}{3}$

C $R = \frac{3A}{2}$

D $R = 6A$

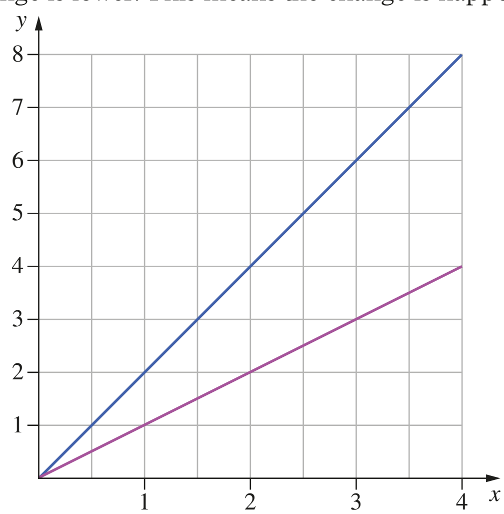
2 Rearrange the formula $y = 3\sqrt{\frac{p}{q}}$ to make p the subject.

6. Interpreting the slope of a linear graph

Scientists often represent a relationship between two variables as a graph. For directly proportional relationships, the variables are connected by a straight line, where the slope (or gradient) of the line represents the constant of proportionality between the two variables.

The slope or gradient of the line is defined as the ratio of change between two points in the vertical axis (ΔY), divided by the change between two points in the horizontal axis (ΔX). In other words, it measures the rate at which one variable (the dependent variable) changes with respect to the other (the independent variable).

The graph below has two straight lines with different slopes. The steeper slope (blue line) indicates that the rate of change is higher. This means the change is happening more quickly. On the other hand, the flatter slope (red line) indicates that the rate of change is lower. This means the change is happening more slowly.



Practice questions

1 On a graph with two sloped lines, what does the steeper sloped line indicate?
A a faster rate of change
B a slower rate of change
C the same rate of change
D a much slower rate of change

2 The rate of change of a straight line on a graph is given by the:
A *y*-intercept
B *x*-intercept
C gradient
D area under the graph

7. Understanding mathematical symbols

Part of the language of science is using symbols to represent quantities or to give meanings. For example, the four symbols <, >, ≤ and ≥ are known as 'inequalities'.

The following mathematical symbols are commonly used in science.

Symbol	Meaning	Example	Explanation
<	less than	$2 < 3$	2 is less than 3
>	greater than	$6 > 1$	6 is greater than 1
≤	less than or equal to	$2x \leq 10$	$2x$ is less than or equal to 10
≥	greater than or equal to	$3y \geq 12$	$3y$ is greater than or equal to 12
$\sqrt{\ }$	square root	$\sqrt{4} = 2$	The square root of 4 is 2
Δ	change in (difference between)	Δt	change in t (time)
≈	approximately equal to	$\pi \approx 3.14$	π is approximately equal to 3.14
Σ	summation	$\sum_{i=1}^{4} i$	The sum of consecutive integers from 1 to 4, i.e. $1 + 2 + 3 + 4 = 10$

Practice questions

1 The symbol that means 'less than' is:
A <
B >
C ≤
D ≥

2 Which of these symbols is an inequality?
A ≈
B Δ
C $\sqrt{\ }$
D ≤

8. Understanding the difference between discrete and continuous data

Quantitative data forms the backbone of science. Scientists are constantly working with data—measuring, recording, analysing and interpreting it.

Quantitative data consists of numerical values that can either be discrete or continuous.

Discrete data is data that has a set of clearly defined values. For example, the number of students in a class would have a discrete set of possible data.

Continuous data is usually data that is measured in some way and can have an infinite number of values. For example, your height or weight would have a continuous set of possible data.

The easiest way to distinguish between the two types of quantitative data is to ask, 'Is the data measured or counted?' If it is counted, the data set is discrete. If it is measured, the data set is continuous.

Practice questions

1 Which one of these data sets is continuous?
- **A** the number of cars parked in a street
- **B** the temperature of the air over a 24-hour period
- **C** the number of students at a school
- **D** the number of nails used to build a fence

2 Which one of these data sets is discrete?
- **A** the number of cars parked in a street
- **B** the temperature of the air over 24 hours
- **C** the heights of a team of footballers
- **D** the mass of a team of netballers

9. Calculating the mean, median and range of a data set

When handling data, scientists often look for ways to describe patterns in the data. Common terms used when analysing a set of data include the mean, median and range.

Mean: the *average* value in the data set. To calculate the mean, sum all the values in the data set and then divide this by the number of data values.

Median: the *middle* value in an ordered data set. To calculate the median, arrange the data set in ascending order and then count the number of data values. If the number of values is odd, the median is the middle value. If the number of values is even, calculate the median by adding the two middle values and dividing by 2, i.e. by calculating the average of the two middle numbers.

Range: the *spread* of values in the data set. To calculate the range, take the largest data value and then subtract the smallest data value.

Practice questions

1 The following set of data is recorded:
44, 17, 21, 26, 42, 18
Find the
- **a** mean
- **b** median
- **c** range.

2 The mass in kilograms of each student in a class of 25 students is recorded below. The combined mass of all the students is 1340 kg.
Find the
- **a** mean
- **b** median
- **c** range.

Students' weights: 67, 60, 41, 52, 39, 60, 42, 55, 55, 50, 46, 62, 48, 48, 56, 64, 55, 56, 59, 61, 41, 63, 53, 62, 45

10. Solving simple algebraic equations

To solve an equation means to find the values that make the equation true. Scientists manipulate equations and substitute in known variables in order to solve for the variable required.

For example, you can solve

$$a = \frac{F_{\text{net}}}{m}$$

where F_{net} is the net force on the car, which is 2400 N
m is the mass of the car, which is 1200 kg
a is the acceleration of the car in $m\,s^{-2}$, which is unknown.

$$a = \frac{2400}{1200}$$
$$= 2\,m\,s^{-2}$$

Practice questions

1 Solve the equation $V = IR$ if $I = 3$ and $R = 9$.
A $V = 3$
B $V = 6$
C $V = 12$
D $V = 27$

2 Solve the equation and find the value of Q if $Q = mc\Delta T$, $m = 1.2$, $c = 4200$ and $\Delta T = 30$.
A $Q = 840$
B $Q = 25\ 200$
C $Q = 126\ 000$
D $Q = 6$

11. Completing calculations with more than one operation

Scientists often deal with complex calculations that can involve numerous operations within the one calculation. The order in which these operations are performed can affect the result of the calculation.

For example, the calculation $2 + 3 \times 4$ will give an incorrect answer of 20 if you calculate the $2 + 3$ part first, but it will give the correct answer of 14 if you calculate the 3×4 part first.

Scientists and mathematicians have agreed on a set order in which operations are carried out so that calculations are consistent. You can remember this order using the acronym 'BIDMAS':

- brackets
- indices (powers, square roots, etc.)
- division and multiplication
- addition and subtraction.

The operations present in a calculation are performed in the order shown in the list. If there are multiple instances of division and multiplication, or addition and subtraction, work from left to right.

For example, the 3×4 part in the original example would always be performed first, since multiplication is higher in the list than addition.

When dealing with scientific notation, it is important to keep each individual number complete. Use brackets to help do this, especially when dividing.

For example, when dividing 3.01×10^{21} by 6.02×10^{23}, brackets are used to keep the second number together.

If you used a calculator and entered $\frac{3.01 \times 10^{21}}{6.02 \times 10^{23}}$ (i.e. $3.01 \times 10^{21} \div 6.02 \times 10^{23}$), the answer would come out as 5.00×10^{43}. This is not the correct answer.

The correct answer is only obtained by entering $\frac{3.01 \times 10^{21}}{(6.02 \times 10^{23})}$. This time, the answer comes out correctly as 5.00×10^{-3}.

Using the calculator's *EXP* button or the $x10^x$ button keeps the number and power of 10 together as one number and avoids the problems of using the number multiplied by 10^x. If your calculator does not have an *EXP* or $x10^x$ button, check the user manual. It may be that it is just labelled differently on your calculator. Alternatively, remember to always use brackets to keep the terms in the denominator together, as shown above.

Practice questions

1 What is 3.4×10^{-4} divided by 1.7×10^{-3}?

A 2×10^{-7}
B 2
C 20
D 0.2

2 Substitute $m = 1.4$, $d = 3.9$ and $c = 2.7$ into $W = 6m - 4(d + c)$ and solve for W.

A –4.5
B –18
C 34.8
D 26.7

12. Understanding the relationship between data, graphs and algebraic rules

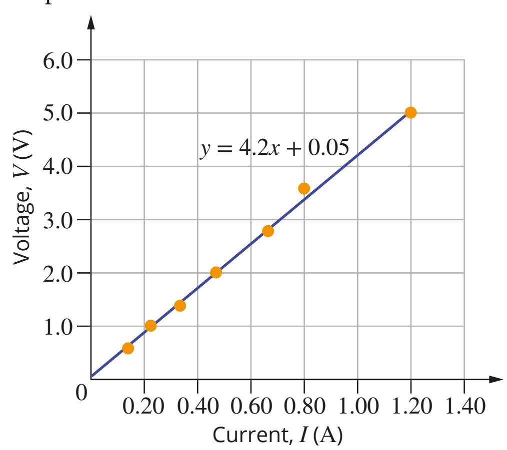
Scientists use graphs to analyse the data they collect from experiments. All graphs tell a story. The shape of the graph shows the relationship between the variables, and this relationship can be written algebraically and numerically. The horizontal axis is known as the x-axis and the vertical axis is known as the y-axis.

Once the algebraic rule is known, the values for one variable can be substituted and the values for the other variable can be calculated. These values can also be determined by reading them from the graph.

For example, when investigating how current and voltage vary across a light bulb, the following data was collected:

Current, *I* (A)	Voltage, *V* (V)
0.14	0.6
0.22	1.0
0.33	1.4
0.47	2.0
0.66	2.8
0.80	3.6
1.20	5.0

Graphing this data produced:



The numerical values from the experiment are listed in the table and plotted on the graph. The algebraic relationship between the variables is given by the equation of the line:

$$y = 4.2x + 0.05$$

The value of the y-intercept is approximately zero, so assuming that the y-intercept is zero, and labelling the x-axis as current and the y-axis as voltage, the relationship can be written as:

$$y = 4.2 \times \text{current}$$

Using the appropriate symbols this can also be written as: $V = 4.2I$

Practice questions

1 If a graph had L on the y-axis and B on the x-axis and the equation of the straight line was $y = 3.7x$, what is the algebraic form of the graph?
A $L = 3.7x$
B $y = 3.7x$
C $y = 3.7B$
D $L = 3.7B$

2 If a relationship was written as $m = 5.9L$, what shape would the graph be and which variable would be plotted on which axis?
A The graph would be non-linear, with m on the y-axis and L on the x-axis.
B The graph would be non-linear, with m on the x-axis and L on the y-axis.
C The graph would be linear, with m on the y-axis and L on the x-axis.
D The graph would be linear, with m on the x-axis and L on the y-axis.

13. Recognising and using ratios

A ratio is the relationship between two numbers of the same kind. It could be the quantities in a recipe, the division of profits from a sale, or the number of different types of the same thing.

Scientists use ratios to compare quantities. This might be the numbers of atoms of different elements in a compound, or the number of primary and secondary windings of a transformer.

You can also use the principle of ratios to solve problems. For example, if 1 reaction produces 2.5 MeV of energy, then how much energy does 34 reactions produce?

The reaction-to-energy ratio of 1:2.5 should remain constant as the number of reactions increases. So you need to find the factor that 1 needs to be multiplied by to give 34:

$$1 \times 34 = 34$$

You then multiply the energy amount by the same factor:

$$2.5 \times 34 = 85\,\text{MeV}$$

You may also see ratios expressed as fractions.

Practice questions

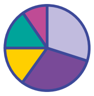
1 If the primary coil of a transformer has 2400 windings and the secondary coil has 600 windings, what is the simplest ratio of primary to secondary windings?
A 2400:600
B 24:6
C 12:3
D 4:1

14. Understanding pie charts, frequency graphs, and histograms

It is essential in science to collect data and arrange it in an orderly way. Tables are often used to organise data, which can then be displayed in a graph.

Pie charts

A pie chart is a circle that is divided into sectors. Each sector represents one item in the data set and is shown as a percentage or fraction of the total data set.



Frequency graphs and histograms

Frequency graphs and histograms are another way of representing data visually.

If data is discrete (i.e. can be counted), each column in a column graph will represent one category, e.g. 'apples' or 'strawberries'. Often these columns have a gap between them.

If the data is continuous (i.e. can be measured), such as the heights of the students in that class, each column will represent a range of possible heights, e.g. 140 to 160 cm, and there will be no gaps between the columns. These are called histograms.

Practice questions

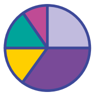
1 Choose the pie chart that correctly shows the data from Emmanuel's poll of soccer fans.

Number of matches watched	10	12	15	18	21
Relative frequency (%)	25	35	15	15	10

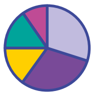
Matches watched:

- = 10
- = 18
- = 12
- = 21
- = 15

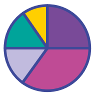
A



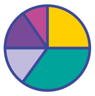
B



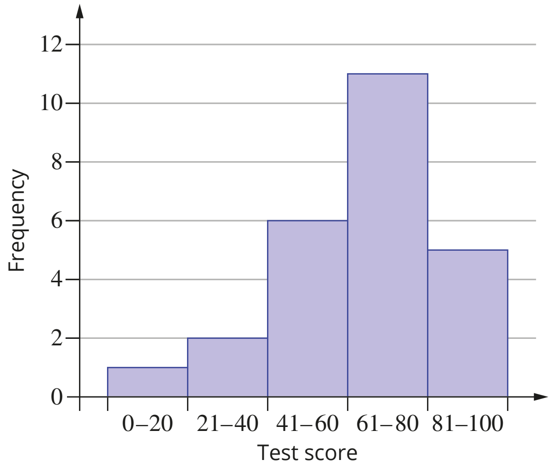
C



D



2 A class of Year 11 Chemistry students had a test. A histogram of their test scores is shown below.



a How many students scored over 80?
b How many students scored below 40?
c How many student are in this class?

15. Understanding the graphical representation of a sine curve

The sine curve is a mathematical curve that describes a smooth repetitive oscillation. It is relevant to the physics topics of sound, AC electricity, simple harmonic motion, waves and many others.

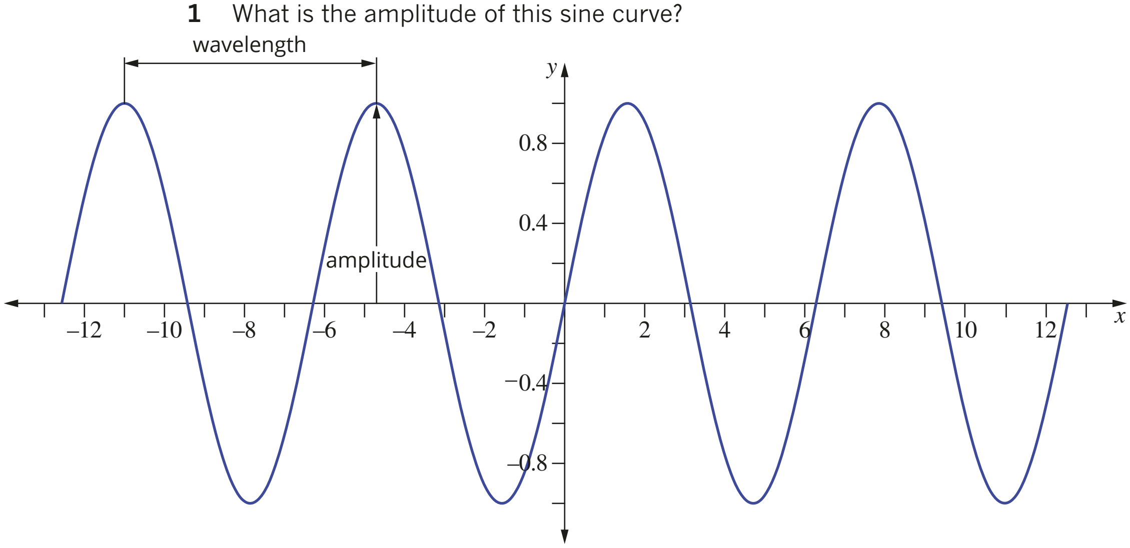
The amplitude of the curve is the distance from the midpoint of the curve to the highest peak or lowest trough, or half the distance from the lowest to the highest point.

The wavelength of the curve is the length of one complete wave.

The period of the curve is the time measurement for one complete cycle of the wave.

Practice questions

1 What is the amplitude of this sine curve?



A 1
B 2
C 4
D 8

2 Sine curves in science can be applied to:
A AC electricity
B simple harmonic motion
C sound waves
D all of the above

16. Understanding sine, cosine and tangent relationships in right-angled triangles

Vector problems in physics (and other applications) often involve using trigonometric relationships to find the unknown side of a right-angled triangle.

You will recall that:

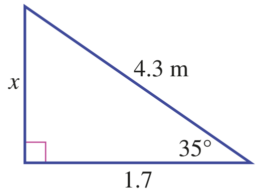
$$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}},\ \cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}},\ \tan\theta = \frac{\text{opposite}}{\text{adjacent}}$$

The acronyms for each of these rules are:

Equations	$\sin\theta$	$\cos\theta$	$\tan\theta$
Acronym	SOH	CAH	TOA

To solve a trigonometry problem, the appropriate formula needs to be identified and solved for the unknown side or angle of a triangle.

In the triangle below, side x is opposite the angle shown as 35°. The hypotenuse, which is always the side opposite the right angle, is 4.3 m long.



θ and the length of the hypotenuse are known. To find the length of the opposite side, use the SOH part of the acronym:

$$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}}$$

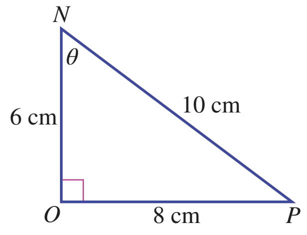
$$\sin 35° = \frac{x}{4.3}$$

$$x = 4.3 \times \sin 35°$$

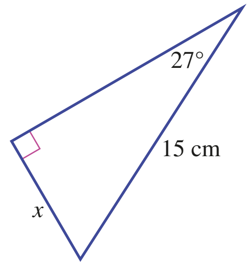
$$= 2.5$$

Practice questions

1 On the following right-angled triangle, label the sides as opposite (O), adjacent (A) and hypotenuse (H) in relation to the angle θ.



2 What is the length of side x in the following triangle? Give your answer to 2 decimal places.



A 6.81 cm
B 13.37 cm
C 7.64 cm
D 12.25 cm

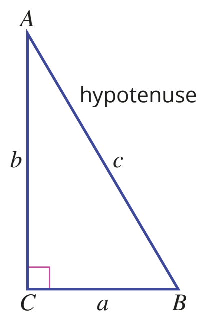
17. Understanding Pythagoras' theorem and similar triangles

Vector problems in physics (and other applications) often involve finding the side lengths of a right-angled triangle and other geometrical unknowns.

Pythagoras' theorem

For any right-angled triangle, Pythagoras' theorem states that the square of the length of the hypotenuse is equal to the sum of the squares of the lengths of the two shorter sides.

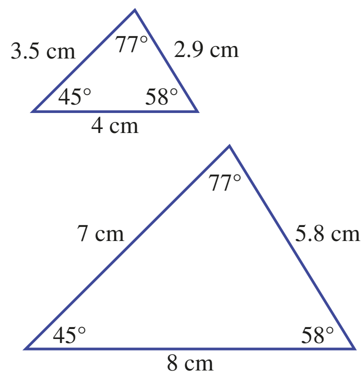
$a^2 + b^2 = c^2$



Similar figures

Similar figures are figures that have the same shape, but are not necessarily the same size.

In similar triangles, all pairs of corresponding angles are equal and all pairs of matching sides are in the same ratio. For example, the triangles shown below are similar triangles.



The corresponding angles in these triangles are equal. The corresponding sides in these triangles are in the same ratio. The ratio can be shown as:

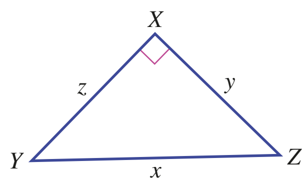
$$\frac{PQ}{AB} = \frac{5.8}{2.9} = 2$$

$$\frac{QR}{BC} = \frac{8}{4} = 2$$

$$\frac{PR}{AC} = \frac{7}{3.5} = 2$$

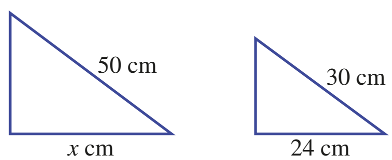
Practice questions

1 For the following triangle, select the correct statement of Pythagoras' theorem.



A $z^2 + x^2 = y^2$
B $z^2 - y^2 = x^2$
C $x^2 + y^2 = z^2$
D $z^2 + y^2 = x^2$

2 These two triangles are similar. Find the value of *x*.



18. Using units in an equation to check for dimensional consistency

Scientists know that each term in an equation represents a quantity. The units used to measure that quantity are not used in the calculations. Units are only indicated on the final line of the solved equation.

For example, this is the equation for the area (A) of a rectangle of length (L) and width (W):

$$A = L \times W$$

If L has a value of 7 m and W has a value of 4 m, it is written:

$$\begin{aligned} A &= L \times W \\ &= 7 \times 4 \\ &= 28\,\text{m}^2 \end{aligned}$$

Note that the units are left out of the actual calculation on the second line and only included at the end, after the numerical answer.

You can use units to check the dimensional consistency of the answer. In the example above, the two quantities of L (length) and W (width) both have to be expressed in consistent units, in this case metres (m), to give an answer that is expressed in square metres ($\text{m} \times \text{m} = \text{m}^2$).

If you had made a mistake, and used the formula $A = L + W$ instead, the answer would be expressed in metres only. This is not the correct unit to express area, so you would know that was wrong.

Practice questions

1 Which formula has the correct dimensions for calculating volume in m^3?
A m × m
B m × s
C m × m × s
D m × m × m

2 Which of these shows the correct substitution into $p = 2L + 2W$ using consistent units for $L = 3.5$ cm and $W = 240$ cm?
A $P = 2 \times 3.5 + 2 \times 240$
B $P = 2 \times 3.5 + 2 \times 24$
C $P = 2 \times 35 + 2 \times 240$
D $P = 2 \times 3.5 + 2 \times 2.4$

3 By using the equation $p = mv$ as a guide, select the correct units in which to measure momentum, p.
A m s
B $m\,s^{-1}$
C $kg\,m\,s^{-1}$
D $kg\,m\,s^{-2}$

19. Understanding inverse and inverse square relationships

Some relationships in science involve one quantity in a relationship increasing and the other quantity decreasing proportionally. This is an inverse relationship.

Thicker wires have lower resistance. This means that, as the cross-sectional area of a wire increases, the value of its resistance decreases. This is an inverse relationship and it can be written as:

$$R \propto \frac{1}{A} \text{ or } R \times A$$

(This shows it is a fixed amount that doesn't change even if R and A change.)

An inverse square relationship is similar, but one quantity increases as the square of the other quantity decreases. Wires with a larger radius will have a lower resistance following an inverse square law.

$$R \propto \frac{1}{r^2} \text{ or } R \times r^2$$

Practice questions

1 Which of these rules represents an inverse relationship?
A $B \propto \frac{1}{L}$
B $y \propto d$
C $X \propto \frac{1}{c^2}$
D $K \propto t^2$

2 Which of these rules represents an inverse square relationship?
A $B \propto \frac{1}{L}$
B $y \propto d$
C $X \propto \frac{1}{c^2}$
D $K \propto t^2$

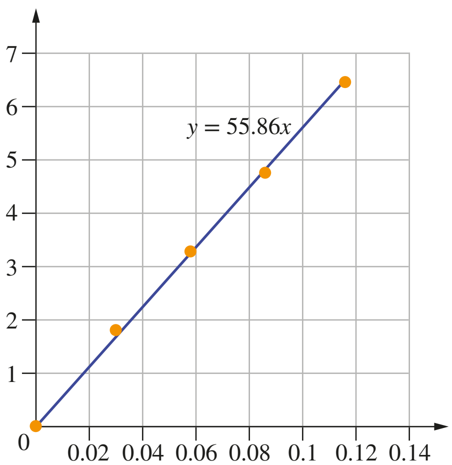
20. Understanding lines of best fit

When scientists observe that the points on a graph seem to form a straight line, then a line of best fit (or trendline) can be drawn through them. Computer programs such as Excel can fit a trendline to a data set; lines of best fit can also be drawn by hand onto a printed or drawn graph.

The line of best fit should pass as close as possible to as many of the points as possible (i.e. it should 'fit' the data closely). It may not pass exactly through any of the points, but once the line of best fit is drawn, the points should be spaced equally on each side, above and below the line. There should be no points very far away from the line, unless they are considered to be unreliable. Unreliable points are called 'outliers' and can be disregarded for the purposes of creating a line of best fit.

The gradient and y-intercept of the line (not the points) can be determined to find the relationship between the variables using the general equation for a straight line: $y = mx + c$.

For example:



Practice questions

1 Are the following statements true or false?

- **a** All the points on a scatter plot must lie on the line of best fit.
- **b** The line of best fit may pass through none of the points.
- **c** The points of the scatter plot should lie close to the line of best fit.

2 Are the following statements true or false?

- **a** A trendline is the same as a line of best fit.
- **b** A line of best fit can only be drawn using a computer program such as Excel.
- **c** Outliers should be included as normal points when considering where to draw a line of best fit.

Answers

Appendix C Maths Skills

1. Transforming decimal notation to scientific notation

1

0.002	2×10^{-3}
2000	2×10^{3}
0.1234	1.234×10^{-1}
12.34	1.234×10^{1}
123.4	1.234×10^{2}

0.002: Move the decimal point 3 places to the right which gives an index of –3, so 0.002 is written as 2×10^{-3}.
2000: Move the decimal point 3 places to the left which gives an index of +3, so 2000 is written as 2×10^{3}.
0.1234: Move the decimal point 1 place to the right which gives an index of –1, so 0.1234 is written as 1.234×10^{-1}.
12.34: Move the decimal point 1 place to the left which gives an index of +1, so 12.34 is written as 1.234×10^{1}.
123.4: Move the decimal point 2 places to the left which gives an index of +2, so 123.4 is written as 1.234×10^{2}.

2 The –4 index shows the decimal point has been moved 4 places to the right. Move it back 4 places to the left to give 0.000 7009.

2. Identifying significant figures

1 B 0.000 40

2 A 1.66
When two numbers are multiplied, use the smallest number of significant figures in the initial values to give your answer. In George's multiplication, the answer is 1.66408, but as 1.22 has 3 significant figures, the correct answer is 1.66.

3 C 41.00
The zeroes to the right of the decimal point are significant. When writing an answer to a correct number of significant figures, you may need to use rounding if the initial answer has more figures (digits) than you need. These extra figures are called 'non-significant figures.' If the first non-significant figure is ≥5, you round up. If the first non-significant figure is <5, then do not round up. For example, if the initial answer for a calculation was 2.1259 but you only needed an answer to 3 significant figures, you would round it to 2.13. If the initial answer was 2.1241, then to 3 significant figures this becomes 2.12.

4 **a** B 78.2
When 65.3 is added to 12.92, the answer is 78.22, but only if it is assumed that 65.3 is actually 65.30. When adding or subtracting with significant figures, use the smallest number of significant figures provided in the initial values. As there is no way of knowing the accuracy of beyond 65.3, the answer should only be given to 3 significant figures, which is 1 decimal place, 78.2.

b When multiplying or dividing, use the smallest number of significant figures in the initial values to give your answer.

$$\begin{aligned} F_{\text{net}} &= ma \\ &= 78.2 \times 1.250 \\ &= 97.8 \end{aligned}$$

Since the mass value only has 3 significant figures, the answer should also only have 3 significant figures. Therefore, even though the calculated answer is 97.75, the answer to 3 significant figures is 97.8 m s^{-1}.

3. Calculating percentages

1 C 75%

$$\frac{9}{12} = \frac{x}{100}$$
$$\begin{aligned} x &= \frac{9}{12} \times 100 \\ &= 75\% \end{aligned}$$

2 B 6

$$\begin{aligned} x &= \frac{25}{100} \times 24 \\ &= 0.25 \times 24 \\ &= 6 \end{aligned}$$

3 C 12.5%

$$\frac{15}{120} = \frac{x}{100}$$
$$\begin{aligned} x &= \frac{15}{120} \times 100 \\ &= 12.5\% \end{aligned}$$

4. Converting between percentages and fractions

1 B 20%
To write a fraction as a percentage, multiply the fraction by 100.

$$\begin{aligned} \frac{1}{5} &= 0.2 \times 100 \\ &= 20\% \end{aligned}$$

2

0.2%	$\frac{1}{500}$
2.5%	$\frac{1}{40}$
17.5%	$\frac{7}{40}$
35%	$\frac{7}{20}$
44.4%	$\frac{111}{250}$

To change a percentage to a fraction, divide by 100.

$$0.2\% = \frac{0.2}{100} = \frac{2}{1000} = \frac{1}{500}$$
$$2.5\% = \frac{2.5}{100} = \frac{25}{1000} = \frac{1}{40}$$
$$17.5\% = \frac{17.5}{100} = \frac{175}{1000} = \frac{35}{200} = \frac{7}{40}$$
$$35\% = \frac{35}{100} = \frac{7}{20}$$
$$44.4\% = \frac{44.4}{100} = \frac{444}{1000} = \frac{222}{500} = \frac{111}{250}$$

5. Changing the subject of an equation

1 C

$R = \frac{3A}{2}$

Multiply both sides of the equation by 3

$3A = 3 \times \frac{2}{3}R$

$= 2R$

Divide both sides of the equation by 2

$\frac{3A}{2} = R$

Rewrite the equation so it reads

$R = \frac{3A}{2}$

2 $p = \frac{y^2 q}{9}$

Divide both sides of the equation by 3

$\frac{y}{3} = \sqrt{\frac{p}{q}}$

Square both sides

$\left(\frac{y}{3}\right)^2 = \frac{p}{q}$

Expand the brackets on the left

$\frac{y^2}{9} = \frac{p}{q}$

Multiply both sides by q

$\frac{y^2 q}{9} = p$

Rewrite the equation so it reads

$p = \frac{y^2 q}{9}$

6. Interpreting the slope of a linear graph

1 A. The steepness of the slope indicates the rate of change. A line with a steeper slope indicates a faster rate of change.

2 C. The gradient or slope of a linear graph indicates the rate of change.

7. Understanding mathematical symbols

1 A <. One way is to remember that the smaller end of the shape points toward the smaller number. For example, $3 < 6$ means 3 is less than 6.

2 D ≤. The symbols <, >, ≤ and ≥ are all inequalities.

8. Understanding the difference between discrete and continuous data

1 B the temperature of the air over a 24-hour period.
If the data can be counted, the data set is discrete. If the data can be measured, the data set is continuous. The temperature of the air can be measured with a thermometer, so the data set is continuous.

2 A the number of cars parked in a street.
If the data can be counted, the data set is discrete. If the data can be measured, the data set is continuous. The number of cars parked in a street can be counted, so the data set is discrete.

9. Calculating the mean, median and range of a data set

1 **a** $44 + 17 + 21 + 26 + 42 + 18 = 168$

$\frac{168}{6} = 28$

b Place the numbers in ascending order: 17, 18, 21, 26, 42, 44
As there is an even number of values, add the two middle values and divide by 2

$\frac{21+26}{2} = 23.5$

c $44 - 17 = 27$

2 **a** $\frac{1340}{25} = 53.6$ kg

b 55 kg

c $67 - 39 = 28$ kg

10. Solving simple algebraic equations

1 D $V = 27$
Substitute the values and solve the equation.
$V = 3 \times 9$
$= 27$

2 B $Q = 25\,200$
Substitute the values and solve the equation.
$Q = 0.2 \times 4200 \times 30$
$= 25\,200$

11. Completing calculations with more than one operation

1 D 0.2
The correct calculation, using brackets, is:

$\frac{3.4 \times 10^{-4}}{(1.7 \times 10^{-3})} = 0.2$

2 B −18

$W = 6 \times 1.4 - 4 \times (3.9 + 2.7)$
$= 6 \times 1.4 - 4 \times 6.6$
$= 8.4 - 26.4$
$= -18$

12. Understanding the relationship between data, graphs and algebraic rules

1 D. $L = 3.7B$
Substituting L for y and B for x gives $L = 3.7B$.

2 C. The graph would be linear, with m on the y-axis and L on the x-axis.
The graph would be linear as the equation is written in the form $y = mx + 0$ with m on the y-axis and L on the x-axis.

13. Recognising and using ratios

1 D 4:1
The primary to secondary ratio is 2400:600. Dividing both sides of the ratio by 600 simplifies it to 4:1.

14. Understanding pie charts, frequency graphs, and histograms

1 A

The largest sector of the pie chart is purple, representing the percentage of people who watched 12 matches (35%). The next in size is blue (10 matches, 25%), then green (18 matches, 15%) and yellow (15 matches, 15%). The smallest sector is red (21 matches, 10%).

2 **a** 5

b 3

c 25
The height of the 81–100 column is 5. This shows 5 students scored over 80.
The heights of the 0–20 and 21–40 columns are 1 and 2. This shows that 3 students scored below 40.
If you add up the heights of all the columns ($1 + 2 + 6 + 11 + 5 = 25$), it tells you that there are 25 students in the class.

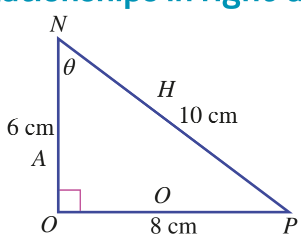
15. Understanding the graphical representation of a sine curve

1 A 1. The amplitude of the curve is the distance from the midpoint (0) to the highest peak (1) or the lowest trough (–1). You could also work out distance from the lowest point to the highest point (1 + 1 = 2) and halve that figure. The amplitude of this sine curve is 1.

2 D all of the above. Sine curves have many applications in physics including those mentioned here.

16. Understanding sine, cosine and tangent relationships in right-angled triangles

1



The hypotenuse is the longest side and the one opposite the right angle, so the 10 cm side is labelled *H*.
The opposite side is the side opposite the θ, so the 8 cm side is labelled O.
The adjacent side is the side next to the angle θ, so the 6 cm side is labelled A.

2 A. 6.81 cm
First, identify the correct trigonometric formula to use. You know the angle 27°, and the hypotenuse is 15 cm long. You want to find the length of *x*, the side opposite the angle, so use the SOH or sine formula.

$$\sin\theta = \frac{\text{opposite}}{\text{hypotenus}}$$

$$\sin 27° = \frac{x}{15}$$

$\sin 27° \times 15 = x$
$x = 6.81\,\text{cm}$
The side labelled x is 6.81 cm long.

17. Understanding Pythagoras' theorem and similar triangles

1 D $z^2 + y^2 = x^2$
The hypotenuse, which is the longest side of a right-angled triangle, is opposite the right angle. In this case it is side x.
x^2 is the sum of the squares of the other two sides: $x^2 = z^2 + y^2$

2 $x = 40\,\text{cm}$
Identify the matching sides. These will be in the same ratio. The hypotenuse of the first triangle is 50 cm. The hypotenuse of the second triangle is 30 cm. The ratio can be written as $\frac{50}{30}$.
The other matching sides will have the same ratio.

$$\frac{x}{24} = \frac{50}{30}$$

$$x = \frac{50 \times 24}{30}$$

$$= 40$$

The value of x is 40.

18. Using units in an equation to check for dimensional consistency

1 D m × m × m

2 D $P = 2 \times 3.5 + 2 \times 2.4$
Both units must be consistent.
Since 100 cm = 1 m the consistent units must be either
$L = 3.5\,\text{m}$ and W = 240 cm
or
$L = 350$ cm and $W = 240\,\text{cm}$
The only correct combination is $P = 2 \times 3.5 + 2 \times 2.4$.

3 C $\text{kg}\,\text{m}\,\text{s}^{-1}$
The unit for momentum is taken from the unit for mass (kg) multiplied by the unit for velocity ($\text{m}\,\text{s}^{-1}$). Therefore, momentum is measured in $\text{kg}\,\text{m}\,\text{s}^{-1}$.

19. Understanding inverse and inverse square relationships

1 A $B \propto \frac{1}{L}$
An inverse relationship contains a term in the form $\frac{1}{x}$.

2 **C** $X \propto \frac{1}{c^2}$
An inverse square relationship contains a term in the form $\frac{1}{x^2}$.

20. Understanding lines of best fit

1 **a** False.
b True.
c True.

2 **a** True.
b False.
c False.

Glossary

A

absolute zero The coldest possible temperature.

absorbed dose The amount of ionising radiation absorbed per kilogram of irradiated material, measured in grays (Gy).

absorption The taking up and storing of energy, such as radiation, light or sound, without it being reflected or transmitted. During absorption the energy may change from one form into another. When radiation strikes the electrons in an atom, the electrons move to a higher orbit or state of excitement by absorption of the radiation's energy.

acceleration The rate of change of velocity. Acceleration is a vector quantity. The SI unit for acceleration is $m\,s^{-2}$.

activity The number of nuclei of a radioactive substance that decay each second, measured in becquerels (Bq).

air pressure The force per unit area exerted by air on an object; related to the density or the number of particles.

air resistance The retarding force (drag) caused by collisions between air and moving objects.

alpha particle A particle consisting of two protons and two neutrons ejected from the nucleus of a radioactive nuclide.

alternating current In an alternating current (AC), electrons oscillate backwards and forwards around a mean position, as opposed to direct current (DC). Household power supplies usually operate at 240 V AC.

amplitude The maximum displacement of a particle from the average or rest position.

angle of incidence The angle an incident ray makes with the normal to the surface it strikes.

angle of reflection The angle a reflected ray makes with the normal to the surface it strikes. Equal to the angle of incidence.

angle of refraction The angle a refracted ray makes with the normal to the surface as it travels from one medium to another.

antineutrino A neutral subatomic particle that interacts very weakly with other matter; the antimatter particle of neutrino.

antinode Areas in a standing wave where complete constructive interference is happening.

artificial transmutation The changing of one element or isotope into another. This happens during radioactive decay and during neutron bombardment in a nuclear reactor.

atomic number The number of protons in a nucleus.

B

background radiation The low level of ionising radiation that exists in the environment as a result of the Earth being radioactive.

beta particle An electron or positron ejected from the nucleus of a radioactive nuclide.

binding energy Energy required to split a nucleus into its separate nucleons.

C

centre of mass A single point in an object where the mass can be considered to be 'concentrated' for the purposes of analysing motion.

chain reaction A series of nuclear fissions that may be controlled or uncontrolled.

charge A property of matter that causes electric effects. Protons have positive charge, electrons have negative charge and neutrons have no charge.

circuit breaker A device that automatically switches off an excessive current by detecting the magnetic field associated with it.

collinear Lying on the same straight line.

components The components of a force are two vectors at right angles to each other that when added together will be equivalent to the original force.

compressions Areas of high pressure in a wave.

conduction The movement of energy (such as heat) from one object to another without the net movement of particles (atoms or molecules).

conductor A substance, body or system that readily conducts heat, electricity, sound or light.

conservation of energy The energy in a system before an interaction is exactly equal to the energy in the system after the interaction.

conservation of mechanical energy The total mechanical energy in a system (i.e. the potential and kinetic energies) remains constant.

conserved When a quantity that exists before an interaction is exactly equal to the quantity that exists after the interaction.

constructive interference The process where two or more waves combine or superpose to reinforce each other. This occurs where the displacement of the individual waves is in the same direction so the amplitude is increased.

contact forces Forces that exist when one object or material is touching another. Friction, drag and normal reaction forces are contact forces.

control rod Material, commonly boron, steel or cadmium, that absorbs neutrons in a nuclear reactor.

controlled variable A variable that must be kept constant during an investigation.

convection A process of heat transfer through a gas or liquid by bulk motion of hotter material into a cooler region.

conventional current A flow of positive electric charge. Conventional current is in the opposite direction to electron flow.

coolant A substance, commonly water, carbon dioxide or liquid sodium, used to transfer thermal energy from the core of a nuclear reactor.

core Part of a nuclear reactor where nuclear fission occurs and thermal energy is produced.

coulomb The SI unit of charge; 1C is equivalent to the combined charge of 6.2×10^{18} protons.

crest The maximum positive displacement reached when particles in a transverse wave are displaced upwards from the average position, or resting position.

critical angle For refraction, this is the incident angle at which total internal reflection occurs. That is the refracted angle is exactly 90 degrees from the normal and lies along the interface between the two media.

critical mass The minimum amount of enriched fissile material in the shape of a sphere that leads to a sustained fission reaction.

current The net flow of electric charge. Current is measured in amperes (A) where $1\ A = 1\ C\ s^{-1}$. By convention, electric current is assumed to flow from positive to negative.

D

daughter nucleus A nucleus on the product side of nuclear equation that results when a nucleus undergoes fission or radioactive decay.

decay series A sequence of radioactive decays that results in the formation of a stable isotope.

dependent variable The variable that may change in response to a change in the independent variable. On a graph, the dependent variable is plotted on the vertical axis.

destructive interference The process in which two or more waves combine or superpose to reduce the amplitude. This occurs where the displacement of the individual waves is in the opposite direction.

deuterium An isotope of hydrogen with one proton and one neutron.

diffract The process affecting light and other wave forms that causes the wave to spread out as the wave passes through a narrow aperture or past an edge.

diffuse Spread out; for example, a wave reflecting off an irregular surface.

dimension Space can be considered to consist of three length dimensions. These length dimensions are arranged at 90 degrees to each other with their point of intersection being the origin. The position of an object can be defined in relation to its position along each of the three dimensions. Typically, these three dimensions are labelled x, y and z. However, up-down, left-right and backward-forward are also appropriate.

dimensional analysis Using the units in a graph or formula to check that the derived term is correct.

direct current In a direct current (DC), electrons travel in one direction only, as opposed to alternating current (AC). Batteries and electric cells provide direct current.

direction conventions Standardised systems for describing the direction in which an object is travelling. The use of cardinal points of a compass (N, S, E and W) is an example of a direction convention.

displacement An object's change in position, relative to its starting position and final position. Displacement does not consider the route the object took to change position, only where it started and where it ended. Displacement is a vector quantity. It is measured in metres (m) and given the symbol s.

distance travelled How far an object travels during a particular motion or journey. Distance is a scalar value. Direction is not required when expressing magnitude. It is measured in metres (m) and given the symbol *d*.

Doppler effect A change in the observed frequency of a wave, such as sound or light that occurs when the source and observer are in motion relative to each other.

dose equivalent A measure of the biological damage inflicted on a tissue due to absorption of a defined quantity of radiation. Dose equivalent measurements take into account the nature of the radiation applied. It is measured in sieverts (Sv).

E

earth The third wire (usually green or green and yellow) in electrical devices that acts as an important safety feature by carrying excess current due to a device malfunction directly into the Earth.

echo The reflection of sound from a distant surface that reaches the ear in more than 0.1 seconds and is therefore heard as a separate sound to the original sound.

effective resistance A single resistance that could be used to replace a number of individual resistors for the purpose of circuit analysis.

efficiency The percentage of energy that is effectively transformed by a system.

elastic collision Collision in which kinetic energy is conserved.

electric current The flow of charged particles.

electric shock Also known as electrocution, in which excess electricity flows into the human body due to a device malfunction or electrical accident.

electrical potential energy Potential energy due to the separation of charge in part of an electric circuit.

electricity A form of energy resulting from the existence of charged particles (electrons or protons). Electricity is fuelled by the attraction of particles with opposite charges and the repulsion of particles with the same charge.

electromagnetic radiation A wide range of frequencies (or wavelengths) that can be created by accelerating charges, which result in a rapidly changing magnetic field and electric field travelling out from the source.

electromagnetic spectrum The entire range of electromagnetic radiation. Consists of radio waves, microwaves, infrared radiation, visible light, ultraviolet, X-rays and gamma rays. In a vacuum, all electromagnetic radiation travels at 3.0×10^8 m s^{-1}.

electron A negatively charged particle in the outer region of an atom; it can move from one object to another, creating an electrostatic charge. When electrons move in a conductor, they constitute an electric current.

electron flow The net flow of electrons. Although electric current is assumed to flow from positive to negative, electrons physically move from negative to positive.

electronvolt (eV) A small unit of energy. One electronvolt (1 eV) is the energy an electron would gain when accelerated across a potential difference of one volt: 1 eV = 1.6×10^{-19} J.

electrostatic force A force that acts between charged particles and can act over relatively large distances.

elementary charge The magnitude of the charge on an electron or proton: $e = 1.6 \times 10^{-19}$ C.

emit Give out. Energy can be emitted in the form of heat, light, radio waves etc.

energy An object possesses energy if it has the ability to do work. Energy takes many forms, for example kinetic energy and potential energy.

evaporation The changing of a liquid into a gas, often under the influence of heat (At a temperature below the boiling point).

F

fast breeder reactor A fast breeder reactor is a nuclear fission reactor in which some neutrons from the fission of uranium-235 are absorbed by non-fissile uranium-238. The non-fissile U-238 usually makes up about 99.3% of nuclear fuel, while the U-235 makes up about 0.7%, however in fast breeder reactors the fuel is enriched until the percentage of fissile U-235 is between 15 and 30%. After absorbing a neutron the U-238 undergoes two beta-minus decays in a relatively short period of time to transmutate into the fissile Plutonium-239 isotope. The Pu-239 can be extracted and used as a fuel for another nuclear fission reactor. The term `fast' in the name refers to the fact that fast neutrons are more effectively absorbed by U-238 than slow neutrons, and so a moderator is not required. This would normally be a problem for normal fission reactors, however at the level of enrichment of the fuel in fast breeder reactors the absorption of fast neutrons by U-235 will still occur at a sufficient rate to sustain both the chain reaction and the plutonium production process.

first harmonic Also known as the fundamental; the longest resonant wavelength in a string or a pipe. For a string or a pipe open at both ends, it is half a wavelength and consists of a node in pressure at each end and an antinode in the middle. For a pipe closed at one end it is a quarter of a wavelength and consists of a node in pressure at the open end and an antinode at the closed end.

fissile Capable of undergoing nuclear fission after capturing low-energy neutrons.

fission When a nucleus splits into two or more pieces, usually after bombardment by neutrons.

fission fragments Nuclides formed during nuclear fission; these are usually radioactive.

forcing frequency The frequency of the force applied to an oscillating substance or object

force A vector quantity which measures the magnitude and direction of a push or a pull. It is measured in newtons (N).

free fall A motion whereby gravity is the only force acting on a body.

frequency A measure of the rate at which something occurs, for example the number of vibrations or cycles that are completed per second or the number of complete waves that pass a given point per second. Measured in hertz (Hz).

fuel rod Long, thin rod of enriched uranium used in a nuclear reactor.

fundamental The lowest and simplest form of vibration, with one antinode.

fuse A circuit device that melts when too much current flows through it, breaking the circuit in the process and protecting the other circuit components.

fusion A process taking place inside stars in which small nuclei are forced together to make larger nuclei. Energy is released in the process.

G

gamma ray High-energy electromagnetic radiation ejected from the nucleus of a radioactive nuclide.

Geiger counter A device for measuring radioactive emissions.

gravitational potential energy Energy available to an object due to its position in a gravitational field. Measured in joules (J).

H

half-life The time taken for half of the nuclei of a radioactive isotope to decay.

harmonic The resonant frequencies produced when standing waves are formed in a string or air column.

heat The energy transferred from a hotter object to a cooler one that increases the kinetic and/or potential energy of the particles in the cooler object.

heat exchanger Part of a nuclear reactor where heat drawn from the reactor core is used to turn water into steam.

heavy water Water that has a higher than normal proportion of water molecules that contain deuterium.

I

impulse The change in momentum of an object is also called the impulse of an object. The impulse is calculated by the final momentum minus the initial momentum.

incident Arriving at or striking a surface, especially a beam of light or radiation, or particles.

independent variable The variable that is selected and deliberately changed by the researcher. On a graph, the independent variable is plotted on the horizontal axis.

inelastic collision Collision in which kinetic energy is not conserved.

inertia A property of an object, related to its mass, that opposes changes in motion.

insulator A material or an object that does not easily allow heat, electricity, light or sound to pass through it. Air, cloth and rubber are good electrical insulators; feathers and wool are good thermal insulators.

intensity A measurement of the energy transmitted by a wave or radiation, given by the square of the amplitude.

internal energy The total kinetic and potential energy of the particles within a substance.

inverse square law Relationship between two variables where one is proportional to the reciprocal of the square of the other.

ion Atom of a chemical element in which the number of electrons and protons is not equal and therefore the atom is electrically charged. If extra electrons are present, the ion has a negative charge. If electrons are missing, the ion has a positive charge.

ionising ability The ability of particles or radiation to ionise matter.

ionising radiation Radiation with enough energy to alter the molecular structure of matter by displacing one or more electrons from an atom and thus creating electrically charged ions.

isotope Atoms with the same number of protons but with different numbers of neutrons.

J

junction A point in an electric circuit from which current can flow into or out of from more than one direction.

K

kelvin An absolute temperature scale based on the triple point of water.

kilowatt hour (kWh) Unit of energy equivalent to 3.6 megajoules. The equivalent amount of energy as a 1000 W device turned on for one hour. It is the unit of measure of electricity usage that is measured by electricity meters and appears on electricity bills.

kinetic energy The energy of a moving body, measured in joules (J).

kinetic particle model A model that states that the small particles (atoms or molecules) that make up all matter have kinetic energy, which means that all particles are in constant motion, even in solids.

L

latent heat The 'hidden' energy used to change the state of a substance at the same temperature, i.e. the energy is not seen as a change in temperature.

latent heat of fusion The energy required to change 1 kg of solid to a liquid at its melting point.

latent heat of vaporisation The energy required to change 1 kg of liquid to a gas at its boiling point.

longitudinal A longitudinal wave is one in which the vibration of the particles within the medium are parallel to the direction of energy flow of the wave.

M

magnitude The size or extent of something, with no need for direction. In physics, this is usually a quantitative measure expressed as a number of a standard unit.

mass An amount of matter. One kilogram of mass is equal to the amount comprising the standard kilogram cylinder of platinum–iridium. Mass can be defined by the amount of matter that would result in an acceleration of $1\,m\,s^{-2}$ when a force of 1 N is applied in a frictionless environment.

mass number The number of nucleons (protons and neutrons) in a nucleus.

mean The average value that is calculated by taking the sum of all values and then dividing by the total number of values.

mechanical energy The energy that a body possesses due to its position or motion. Kinetic energy, gravitational energy and elastic potential energy are all forms of mechanical energy.

mechanical wave A wave that transfers energy through a medium.

median The middle piece of data when a data set is listed in order.

medium The material or substance through which a mechanical wave moves.

metal Material in which some of the electrons are only loosely attracted to their atomic nuclei. The properties of metals include: high strength, good electrical and thermal conductivity, lustre, malleability and ductility.

mode The most common piece of data in a data set.

moderator A material, usually graphite or water, that slows neutrons in a nuclear reactor.

momentum The product of an object's mass and velocity. Objects with larger momentum require a larger force to stop them in the same time that an object with smaller momentum takes to stop. It is given by the equation $p = mv$.

N

natural frequency The specific frequency at which an object will tend to vibrate.

net charge When the number of positive and negative charges in an object is not balanced.

net force The vector sum of all the individual forces acting on a body.

neutral No electric charge, or a situation in which positive and negative charges are balanced.

neutron An uncharged subatomic particle.

newton SI unit of force. One newton (1 N) is defined as the force required to make a mass of 1 kg accelerate at $1\,kg\,m\,s^{-2}$.

Newton's first law States that an object will maintain a constant velocity unless an unbalanced, external force acts on it.

Newton's second law States that force is equal to the rate of change of momentum. This can be processed mathematically to: the acceleration of an object is directly proportional to the force on the object and inversely proportional to the mass of the object.

Newton's third law States that for every action (force), there is an equal and opposite reaction (force).

node Areas in a standing wave where complete destructive interference is occurring and the two waves totally cancel each other out.

non-contact forces Forces that act at a distance and do not require the bodies to actually touch each other. Strong nuclear, weak nuclear, gravitational and electromagnetic forces are non-contact forces.

non-ionising radiation Radiation that does not have enough energy to break the molecular bonds within molecules and to alter the number of electrons in an atom. Lower forms of energy in the electromagnetic spectrum such as radio waves, microwaves, visible light and UVA radiation are non-ionising.

non-metal Material in which all of the electrons are strongly attracted to their atomic nuclei.

non-ohmic Not behaving according to Ohm's law; resistance changes depending on the potential difference.

normal An imaginary line at 90°, i.e. perpendicular, to a surface.

nuclear transmutation The changing of one element into another.

nucleon A particle located in the nucleus of an atom.

nucleus The central part of an atom.

nuclide The range of atomic nuclei associated with a particular atom, which is defined by its atomic number, and the various isotopes of that atom as identified by the mass number.

O

ohmic A resistor that follow's Ohm's law; i.e. has a linear relationship between the current it draws and the potential difference across it.

oscillate The movement of particles about their average position in a regular, repetitive or periodic pattern.

outlier A value that lies outside the main group of data of which it is a part. Outliers in data could be caused by errors in the experiment.

overload When an unsafe amount of current flows through a wire; for example, when too many electrical appliances are connected to the same power point.

overtone A harmonic (resonant frequency) that is higher than the natural frequency.

P

parallel circuit A circuit that contains junctions; the current drawn from the battery, cell or electricity supply splits before it reaches the components and re-joins afterwards.

parent nucleus A nucleus on the reactant side of a nuclear equation that when struck by a neutron undergoes fission or simply decays by natural means.

particle displacement The measure of the distance a particle moves about its equilibrium position during the propagation of a wave. In a longitudinal wave this motion is parallel to the direction of propagation of the wave. In a transverse wave it is perpendicular to the direction of travel.

passive heating Energy efficient design that lead to little or no mechanical heating requirements in a building.

penetrating ability A measure of how easily radiation passes through matter.

period The time interval for one vibration or cycle to be completed.

personal protective equipment (PPE) Equipment such as safety glasses and disposable gloves used to protect people working in the laboratory.

phase When two or more waves of the same wavelength and amplitude exactly line up.

plane wave A wave that has a straight wave front.

position The location of an object with respect to a reference point. Position is a vector quantity.

positron The antimatter pair of the electron. This means it shares the same mass as an electron but has opposite properties like electromagnetic charge and spin.

potential difference The difference in electric potential between two points in a circuit; measured by a voltmeter when placed across a circuit. A battery creates the potential difference across a circuit, which drives the current.

potential energy Energy that can be considered to be 'stored' within the field due to an object's position within the field, composition or molecular arrangement.

power The rate at which work is done; a scalar quantity measured in watts (W).

proton A positively charged subatomic particle.

pulse A single movement, vibration or undulation.

Q

qualitative variable A variable that can be observed but not measured.

quality factor The number used to indicate the weighting of the biological impact of radiation.

quantitative variable A variable that can be measured.

R

radiation Rays or particles that carry energy. Also, the process by which energy is emitted by an object or system, transmitted through an intervening medium or space, and absorbed by another object or system.

radiation shield A thick concrete wall that prevents neutrons escaping from a nuclear reactor.

radioactive Something that spontaneously emits radiation in the form of alpha particles, beta particles and gamma rays.

radioisotope An isotope of a chemical element that emits radioactivity due to its unstable combination of neutrons and protons in the nucleus.

random error An error in measurement that occurs in an unpredictable manner.

rarefaction An area of decreased pressure within a longitudinal sound wave.

raw data The actual measurements taken directly during an investigation without being processed in any way.

ray A line drawn perpendicular to a wave front and in the direction the wave is moving. (Also a narrow beam of light.)

reflection The change of direction of a wave as it strikes a surface and is bounced back.

refraction The bending of the direction of travel of a ray of light, sound or other wave as it enters a medium of differing refractive index (optical density).

reliability The consistency of the results obtained from an experiment or collection of data. Reliable results are also repeatable, meaning another scientist performing the same analysis will come up with the same results.

residual current device (RCD) A device that can detect a difference in the active and neutral wires and switch off current in dangerous situations to help prevent electrocution.

resistance A measure of how much an object or material resists the flow of current; the ratio of the potential difference across a circuit component and the current flowing through it: $R = V/I$. Resistance is measured in ohms (Ω).

resistor A circuit component, often used to control the amount of current in a circuit by providing a constant resistance. Resistors are ohmic conductors, i.e. they obey Ohm's law.

resonance The state of a system in which an abnormally large vibration is produced in response to an external vibration. Resonance occurs when the frequency of the vibration is the same, or nearly the same, as the natural vibration frequency of the system.

resonant frequency The natural frequency at which an object tends to vibrate.

resultant One vector that is the sum of two or more vectors.

reverberation This is a reflection of sound from a nearby surface that reaches the ear in less than 0.1 seconds and combines with the original sound. It often sounds like a longer sound.

S

scalar A physical quantity that is represented by magnitude and units only. Mass, time and speed are examples of scalar quantities.

seismic wave Vibrations within the earth caused by phenomena such as earthquakes, explosions, volcanoes and landslides.

series circuit When circuit components are connected one after another in a continuous loop so that the same current passes through each component.

short circuit The situation in which a good conductor is inadvertently placed across a battery and an excessive current flows, which may cause damage.

significant figures The numbers in a measurement or calculation that convey meaning and precision.

sinusoidal In the shape of a sine wave.

specific heat capacity The amount of energy that must be transferred to change the temperature of 1 kg of material by 1°C or 1K.

speed The ratio of distance travelled to time taken. Speed is a scalar quantity. The SI unit for speed is m s^{-1}.

spontaneous transmutation The changing of one element into another in a natural process involving radioactive decay.

standing wave Also called a stationary wave, the periodic disturbance in a medium resulting from the combination of two waves of equal frequency and intensity travelling in opposite directions.

strong nuclear force A short-range but powerful force of attraction that acts between all the nucleons in the nucleus. The strong nuclear force acts on quarks and binds them together in hadrons. It also acts at larger distances to bind protons and neutrons together within atomic nuclei.

subcritical mass A quantity of fissile material that is too small to sustain a chain reaction.

supercritical mass A quantity of fissile material that is large enough to sustain a chain reaction.

superposition When two or move waves travel in a medium, the resulting wave at any moment is the sum of the displacements associated with the individual waves.

systematic error An error that is consistent and will occur again if the investigation is repeated in the same way.

T

temperature A measure of the average kinetic energy of the particles in a substance. Temperature can be measured in degrees Celsius (°C) or kelvin (K).

thermal equilibrium For two bodies in thermal contact, the point at which the two reach the same temperature and there is no further net transfer of thermal energy.

total internal reflection Occurs when the angle of incidence exceeds the critical angle for refraction. Light or waves are reflected back into the medium; there is no transmission of light.

tracer A radioactive isotope with a short half-life that is injected into a patient or ingested to monitor biological processes in the body.

transfer The conversion of energy from one system to another.

transform To change from one thing to another; for example, to change energy from electrical potential energy to kinetic energy.

transmit To cause light, heat, or sound, etc. to pass through into a medium.

transuranic Elements with atomic numbers greater than uranium ($Z = 92$). All of these elements are unstable and radioactively decay into lighter elements.

transverse Lying or extending across something. The vibrations of a transverse wave are at right angles to the direction of travel of the wave.

travelling wave A wave that travels unimpeded through a medium and is not confined to a given space. Every point on the wave would have maximum displacement at some point in time. Similarly, each point would also have minimum displacement at some point.

tritium An isotope of hydrogen with one proton and two neutrons.

U

uncertainty The description of the range of data obtained; the maximum variance from the mean.

units Properties related to physical measurements. Units can be fundamental like metres (m), seconds (s) or kilograms (kg). Units can also be derived by combining fundamental units; for example metres per second (m s^{-1}).

V

validity The reasonableness of the results received from an experiment or collection of data. Valid results meet all the requirements of the criteria of the scientific method.

variable A factor or condition that can change.

vector A physical quantity that requires magnitude, units and a direction in order to be fully defined. Velocity, acceleration and force are examples of vector quantities.

vector diagram A system of adding vectors where each vector is drawn head-to-tail, with the resultant vector drawn from the tail of the first vector to the head of the last vector.

velocity The ratio of displacement to time taken. Velocity is a vector quantity. The SI unit for velocity is m s^{-1}.

vibration A repeated motion.

volatile Liquids with weak surface bonds that evaporate rapidly.

volt The unit of electrical potential. One volt is equal to one joule of potential energy given to one coulomb of charge in a source of potential difference. The voltage (or the number of volts) is another name for the potential difference.

voltmeter A device used to measure the electrical potential difference between two points in a circuit.

W

wave front The set of points reached by a wave of vibration at the same instant. Wave fronts generally form a continuous line or surface.

wavelength The distance between one peak or crest of a wave of light, heat or other energy and the next corresponding peak or crest (symbol: λ).

weight The force of attraction on a body due to gravity.

work The transfer of energy as a result of the application of a force; measured by multiplying the force and the displacement of its point of application along the line of action. Measured in joules (J).