Physics

For Secondary Schools

Student's Book Form Four

Tanzania Institute of Education

Physics

for Secondary Schools

Student's Book

Form Four

THE UNITED REPUBLIC OF TANZANIA
MINISTRY OF EDUCATION,
SCIENCE AND TECHNOLOGY

Certificate of Approval

No. 745

Title of Publication: Physics for Secondary Schools Student's Book Form Four

Publisher: Tanzania Institute of Education

Author: Tanzania Institute of Education

ISBN: 978-9987-09-393-9

This book was approved by the Ministry of Education, Science and Technology on 2nd March, 2022 as a Textbook for Form Four students in Tanzania Secondary Schools as per 2007 Syllabus.

Dr Lyabwene M. Mtahabwa
Commissioner for Education

Tanzania Institute of Education

Published 2022

ISBN: 978-9987-09-393-9

Tanzania Institute of Education
P. O. Box 35094
Dar es Salaam, Tanzania

Mobile numbers: +255 735 041 168 / +255 735 041 170
E-mail: director.general@tie.go.tz
Website: www.tie.go.tz

Table of Contents

Acknowledgements v
Preface vi

Chapter One: Waves 1
Concept of wave 1
Propagation of waves 5
Behaviour of waves 7
Sound waves 20
Musical sounds 25
Electromagnetic waves 39
Chapter summary 47
Revision exercise 1 48

Chapter Two: Electromagnetism 51
Concept of electromagnetism 51
Magnetic field due to a current carrying conductor 52
Electromagnetic induction 60
Chapter summary 80
Revision exercise 2 81

Chapter Three: Physics of the atom 86
Concept of an atom 86
Natural radioactivity 93
Artificial radioactivity 104
Radiation hazards and safety 105
Nuclear reactions 107
Chapter summary 112
Revision exercise 3 113

Chapter Four: Thermionic emission 115
Concept of thermionic emission 115
Cathode rays 116
X-rays 120
Chapter summary 123
Revision exercise 4 123

Chapter Five: Electronics 125
Concept of electronics 125
Band theory in solid materials 128
Classification of solid materials in terms of conductivity 129
Types of semiconductors 132
Diodes 140
Transistors 146
Electronic amplifiers 150
Chapter summary 153
Revision exercise 5 154

Chapter Six: Elementary astronomy 157
Concept of astronomy 157
Constellations 161
The solar system 167
Gravitational force 175
The Earth and its Moon 177
Chapter summary 183
Revision exercise 6 184

Chapter Seven: Physics of the Earth and its atmosphere 186
The concept of the Earth and its atmosphere 186
The structure and composition of the Earth 187
Earthquakes and volcanoes 191
Structure and composition of the atmosphere 203
The greenhouse effect and global warming 206
Chapter summary 211
Revision exercise 7 212

Answers to numerical questions 214

Appendix 1: Periodic table 216

Glossary 217

Bibliography 223

Index 224

Acknowledgements

The Tanzania Institute of Education (TIE) would like to acknowledge the contributions of all the organisations and individuals who participated in designing and developing this textbook. In particular, TIE wishes to thank the University of Dar-es-Salaam (UDSM), the University of Dodoma (UDOM), Mkwawa University College of Education (MUCE), Tanzania Atomic Energy Commission (TAEC), School Quality Assurance (SQA) Department, Teachers' colleges and secondary schools. Besides, the following categories of individuals are acknowledged:

Writers: Ms Auguster F. Kayombo (TIE), Mr Shaban J. Baya (TIE), Dr Emmanuel D. Sulungu (UDOM), Mr Daudi T. Mazengo (UDOM) & Mr Mkomwa A. Mtiga (Loyola high school)

Editors: Dr Ismael N. Makundi (UDSM), Dr Innocent J. Lugendo (UDSM), Dr Talam E. Kibona (MUCE), Dr Benard S. Mwankemwa (UDOM), Dr Suleiman A. Suleiman (TAEC), Mr Tungu C. Chagu (MOTCO) & Mr Alphonce A. Mbalwa (Marian girls' secondary school)

Designer: Mr Amani J. Kweka

Illustrators: Mr Fikiri A. Msimbe (TIE), Ms Rehema H. Maganga (TIE), Mr Godlove S. Kyando (TIE) & Alama Art and Media Production Co. Ltd

Coordinator: Ms Auguster F. Kayombo

Furthermore, TIE extends its sincere appreciation to the United States Agency for International Development (USAID)-Tanzania for granting permission to use the materials from 2011 Physics for Secondary Schools, Forms 4, (Revised Edition) textbook.

TIE also appreciates the participation of the secondary school teachers and students in the trial phase of the manuscript. Likewise, the Institute would like to thank the Ministry of Education, Science and Technology for facilitating the writing and printing of this textbook.

Dr Aneth A. Komba
Director General
Tanzania Institute of Education

Preface

This textbook, *Physics for Secondary Schools* is written specifically for Form Four students in the United Republic of Tanzania. It is written in accordance with the 2007 Physics Syllabus for Ordinary Secondary Schools, Form I-IV, issued by the then Ministry of Education and Vocational Training.

The book consists of seven chapters, namely Waves, Electromagnetism, Physics of the atom, Thermionic emission, Electronics, Elementary astronomy and the Earth and the atmosphere. Each chapter contains illustrations, activities, tasks and exercises. You are encouraged to do all the activities, tasks and exercises as well as other assignments that your teacher will provide. Doing so will enable you to develop the intended competencies.

Tanzania Institute of Education

Chapter One

Waves

Introduction

In your daily activities, you may have experienced phenomena such as ripples in a water pond, musical sound from a guitar and tremors caused by earthquakes. All these are wave phenomena. Devices such as television, cellphones, radios, microwave ovens and medical diagnostic machines such as X-ray and ultrasound make use of waves in their operations. Generally, waves are important in our daily activities. In this chapter, you will learn about the concept of waves, propagation of waves, behaviour of waves, sound waves, musical sounds and electromagnetic waves. The competencies developed will enable you to explore how waves are involved in various fields including communication and medicine. You will also acquire skills that will enable you to design and repair some musical instruments and other devices that apply wave principles.

Concept of wave

If you drop a stone into a pond of still water, some disturbance will be created on the water surface. The disturbance, in form of ripples, spreads outward in a circular pattern as shown in Figure 1.1.

Figure 1.1: *Water ripples*

This happens because when the stone hits the water surface, the energy of the stone is transferred to the water molecules at the point of contact. The molecules then vibrate up and down, disturbing the neighbouring water molecules in the process. This results into transfer of energy which is observed as the movement of the disturbance outwards. Basically, the water molecules do not move outwards with the disturbance, only the disturbance moves. The movement of the disturbance is called a wave. A wave can also be observed when you shake one end of a rope up and down. Waves produced by dropping a stone into a water pond or by shaking a rope require a medium for their travel. For the case of dropped stone, the medium is water while for the case of the rope, the medium is the rope itself. Generally, the medium can be a solid, liquid or gas.

A wave is a periodically repeating disturbance that travels through a medium from one location to another, without a net movement of the medium or particles of the medium.

Task 1.1

Given a rope, small stone, basin, water, and slinky spring

1. Tie one end of a stretched rope to a fixed object. Hold the loose end of the rope and gently shake it up and down. Observe the wave motion of the rope.
2. Drop a small stone in a basin that is three quarter filled with water. Observe the wave motion in the water.
3. Place the slinky spring on a flat surface. Hold it on one end and allow one of your friends to displace the spring sideways by approximately one centimetre and then release it. Observe the motion.
4. Discuss your observations with your colleagues in class.

Wave terminologies and parameters

Different terminologies and parameters are used to describe waves. These terminologies and the wave parameters can be explained well using the displacement-distance graph shown in Figure 1.2.

1. Crest

It is the point of maximum positive displacement of the medium particles from the equilibrium position.

2. Trough

It is the point of maximum negative displacement of the medium particles from the equilibrium position (still position).

3. Amplitude

The amplitude, represented by A, is the maximum displacement of the medium particles from the equilibrium position. It is the distance from the central line to the top of a crest or to the bottom of a trough. The SI unit of amplitude is metre (m).

4. Wavelength

This is the distance between two successive crests or two adjacent troughs as shown in Figure 1.2. It is also the distance that the wave travels in one complete cycle WXY. The wavelength is represented by the Greek letter lambda, λ and its SI unit is metre (m).

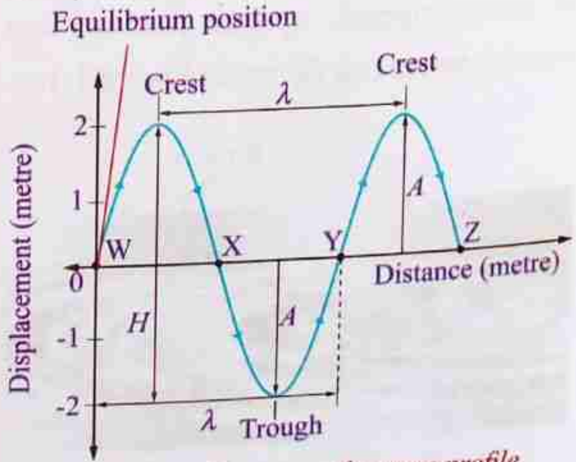


Figure 1.2: *Features of a wave profile*

5. **Period**

This is the time taken for the wave to travel from a crest to the next crest or from a trough to the next trough. It is also the time taken to make one complete cycle WXY as illustrated in Figure 1.3. Period is represented by the letter T and can be calculated by dividing the time of travel, t by the number of complete cycles, n.

That is, $T = \frac{t}{n}$. The SI unit of period is second (s).

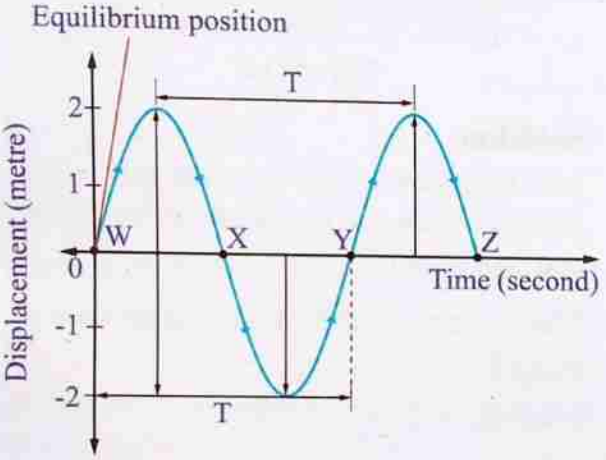


Figure 1.3: *Period of a wave*

6. **Frequency**

The number of waves that pass a certain point per specified amount of time is referred to as *frequency,* represented by the letter f. Frequency can be expressed as the number of wave cycles completed in one second. That is,

$$\text{Frequency, } f = \frac{\text{Number of cycles}}{\text{Time}} = \frac{n}{t},$$

From the definition of period, $t = nT$

Thus, $f = \frac{n}{nT} = \frac{1}{T}$

Hence, $f = \frac{1}{T}$.

This is the relationship between the period, T and frequency, f of a wave. Since T is measured in seconds, then the SI unit of frequency is cycles per second.

For example, if the end of the rope is moved up and down thrice in a second, three waves are produced in this time. Therefore, the frequency of the wave is 3 wave cycles per second. The SI unit of frequency is also known as Hertz (Hz). One Hertz is the same as one cycle per second

7. **Wave velocity**

This is the velocity at which the wave moves through a medium. It is commonly referred to as speed. It is the distance travelled by a wave per unit time. The SI unit of wave velocity is metre per second ($m s^{-1}$).

8. **Phase**

It refers to the angle-like quantity that represents the fraction of a cycle covered in a certain time, t. Phase is normally denoted by 'phi', ϕ. If we consider the points W, X, Y and Z, in Figure 1.3, the velocity of particles corresponding to points W and Y is the same in magnitude and direction. Thus, W and Y are said to be in phase. At points W and X, the particle has the same velocity in magnitude but it takes different directions. Therefore, at these points the wave is out of phase.

Relationship between wave parameters

When a wave completes one cycle, it has travelled a distance, $l = \lambda$. Since the time taken to complete one cycle is T, then wave speed can be determined from the relation,

$$v = \frac{\lambda}{T}$$

But $\quad f = \frac{1}{T}$

Therefore, $v = \lambda f$.

Example 1.1

Determine the amplitude, period and frequency of the wave represented in the displacement-time graph shown in Figure 1.4.

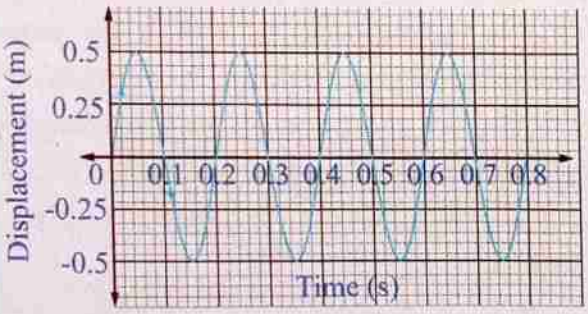


Figure 1.4

Solution

The maximum position is measured from equilibrium 0 to 0.5 m. Therefore, amplitude of the wave is 0.5 m.

The period of the wave, is the time taken to complete one cycle. In Figure 1.4, the time taken to complete one cycle is 0.2 s. Therefore, the period is 0.2 s.

The frequency of the wave is given by:

$$f = \frac{1}{T}$$

$$f = \frac{1}{0.2\ \text{s}}$$

$$f = 5\ \text{Hz}.$$

Example 1.2

The displacement-distance graph corresponding to a wave in example 1.1 is shown in Figure 1.5. Use the graph to determine the wavelength and velocity of the wave.

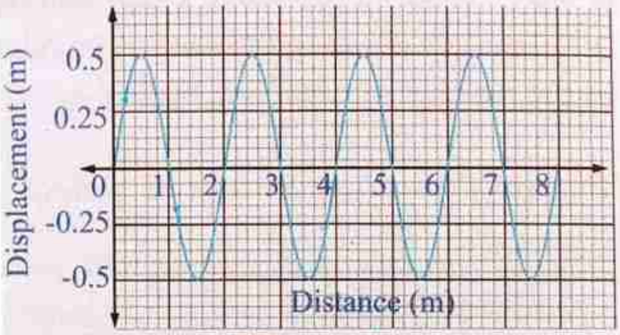


Figure 1.5

Solution

From the graph, the wavelength is the distance between two successive crests and its value is 2 m.

The velocity of the wave, v is given by:

$$v = f\lambda$$

$$v = 5\ \text{Hz} \times 2\ \text{m}$$

$$v = 10\ \text{m s}^{-1}$$

Example 1.3

What is the wavelength of a wave whose speed is 4 m s^{-1} and frequency is 2 Hz?

Solution

Speed, $v = 4\ \text{m s}^{-1}$

Frequency, $f = 2\ \text{Hz} = 2\ \text{s}^{-1}$

Then from,

$$v = f\lambda$$

Thus, $\lambda = \frac{v}{f}$

$$= \frac{4\ \text{m s}^{-1}}{2\ \text{s}^{-1}}$$

$$= 2\ \text{m}.$$

Propagation of waves

The movement of waves in space and time is referred to as wave propagation. Some waves require a medium to propagate, while others do not. Waves that require a material medium to propagate are called mechanical waves. Those waves which do not require material medium to propagate are called electromagnetic waves.

Mechanical waves

A mechanical wave is the wave that is produced when particles vibrate in a medium (solid, liquid or gas) in which the wave propagates. The propagation of a mechanical wave through a medium depends on the elastic and inertia properties of that medium. A typical example of a mechanical wave is sound wave. Other examples, of mechanical waves are water waves and waves on strings. As the mechanical wave propagates, particles of the medium vibrate about their equilibrium positions.

Electromagnetic waves

An electromagnetic wave is a wave that is created as a result of vibrations between an electric field and a magnetic field. The electric and magnetic fields propagate in phase and at right angle to each other. All electromagnetic waves travel through vacuum at the speed of $3\times10^{8}\ \mathrm{m\,s^{-1}}$. Examples of electromagnetic waves include: visible light, radio waves, microwaves, and X-rays.

Modes of wave propagation

All waves can be classified in terms of their modes of propagation. There are two modes of wave propagation; transverse propagation and longitudinal propagation. Waves that propagate in a direction perpendicular to the direction of particle vibrations are called *transverse waves*. On the other hand, waves that propagate in a direction that is parallel to the direction of particle vibrations are referred to as *longitudinal waves*.

Transverse waves

In a transverse wave, the motion of particles make a right angle with the direction of propagation of the wave. For example, when the string or rope under tension oscillates up and down at one end, the disturbance moves along the rope. The particles in the rope vibrate perpendicular to the direction of the disturbance. Figure 1.6 illustrates the propagation of a transverse wave.

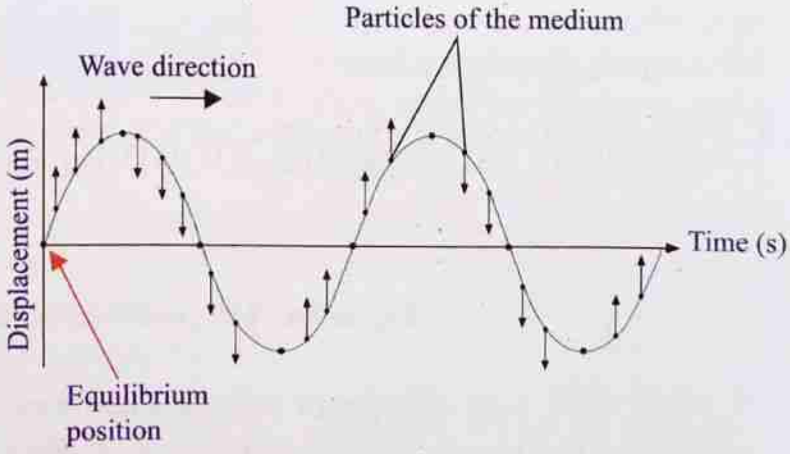


Figure 1.6: *A transverse wave in a string*

Water waves are another example of transverse waves. The water particles move up and down while the waves move in a horizontal direction. That is why a boat on the ocean moves up and down while the waves themselves move towards the shore (Figure 1.7).

Figure 1.7: *Water waves*

Longitudinal waves

In longitudinal waves, particles of a material medium vibrate in a direction parallel to the direction of the wave propagation. A longitudinal wave consists of regions of high and low particle density. These regions are respectively termed as *compression* and *rarefaction regions*. Compression regions are regions of high pressure and high density since particles are being compressed close together. Conversely, rarefaction regions are regions of low pressure and low density as particles are being spread further apart as illustrated in Figure 1.8. The distance between the centres of two consecutive regions of compression or adjacent regions of the rarefaction is the wavelength, λ.

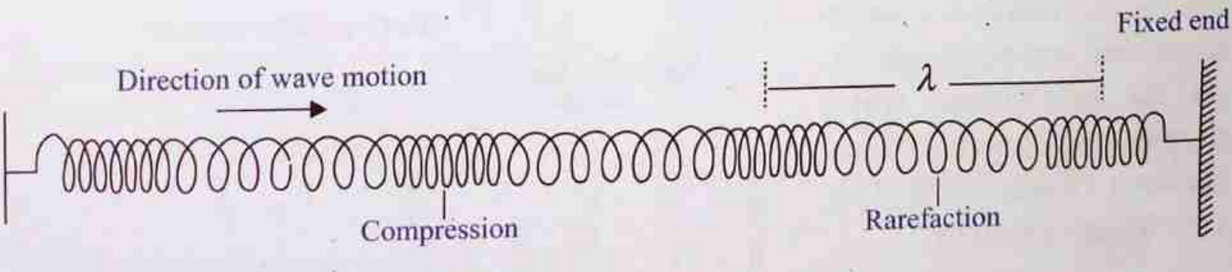


Figure 1.8: *A longitudinal wave in a spring*

A sound wave is an example of a longitudinal wave. It is produced by the vibration of particles and travels through a medium such as air. The amplitude of a sound wave is the difference between the maximum pressure caused by the disturbance and the pressure of the undisturbed particles in a medium as illustrated in Figure 1.9.

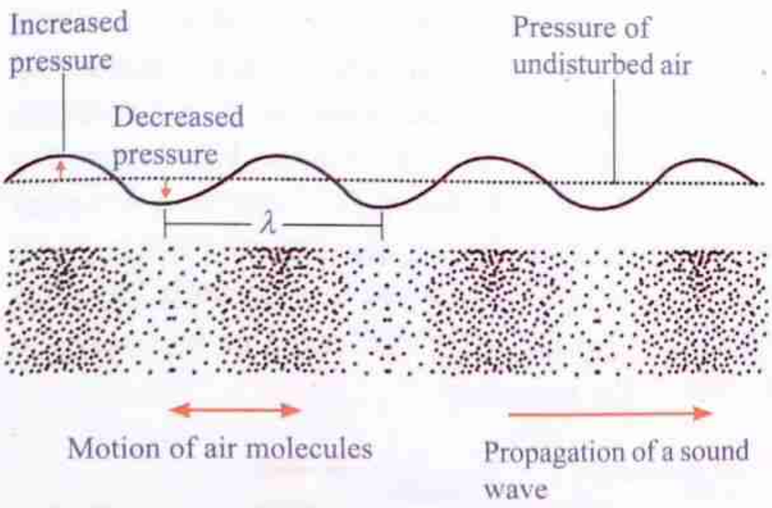


Figure 1.9: *Propagation of sound waves in air*

Note that, in some cases, waves are neither purely transverse nor purely longitudinal. For example, water waves at the surface of a large water body involve components of both longitudinal and transverse waves.

For all electromagnetic waves, oscillations are perpendicular to the direction of wave propagation. Therefore, electromagnetic waves are classified as transverse waves.

Task 1.2

Use your own creativity to create longitudinal and transverse waves and share your experience with your classmates by demonstrating how you create those waves.

Behaviour of waves

In a medium of uniform properties, waves propagate at constant speeds. Any change in medium properties results in a change in the speed of a wave. For example, the propagation speed of sound depends on the type, composition, and temperature of a medium through which it propagates. A change in the speed of the wave means a change in its wavelength, since the frequency remains constant. Thus, in a medium of uniform properties the speed of a wave remains constant so the wavelength varies inversely proportional to the frequency. The amplitude of the wave depends on the amount of energy being transmitted. A decrease in the amplitude of a wave shows that the wave has lost some energy. Change of medium properties result to different behaviours of a waves. These behaviours include reflection, refraction, diffraction and interference.

Reflection of waves

A travelling wave may encounter a boundary between two media of different properties. If the boundary does not allow the wave to pass through, the wave bounces back to the medium in which it was propagating before striking the boundary. This phenomenon is referred to as *reflection*. For example, a wave travelling through a string which is fixed at one end will be reflected upon reaching the fixed end of the string. In this case, the fixed end of the string acts as the boundary.

Consider the situation where a string is fixed to a rigid wall at its right end. This end is called a fixed end. When a wave is allowed to propagate through the string, the wave reaches the fixed end, and gets reflected. The reflected wave will be inverted as shown in Figure 1.10 (a). In this case, the amplitudes of both the incident and reflected waves are the same.

If the right end of the string is tied to a ring, which can slide up and down on a rod without any friction, the end is termed as a free end. In this case, when the wave arrives at the free end, the ring moves up and down. The motion of the free end results into a reflected wave which is not inverted. The reflected wave will have the same speed and wavelength as the incident wave. However, the reflected wave will have a smaller amplitude as shown in Figure 1.10 (b). The decrease in amplitude indicates that the wave lost some of its energy at the boundary.

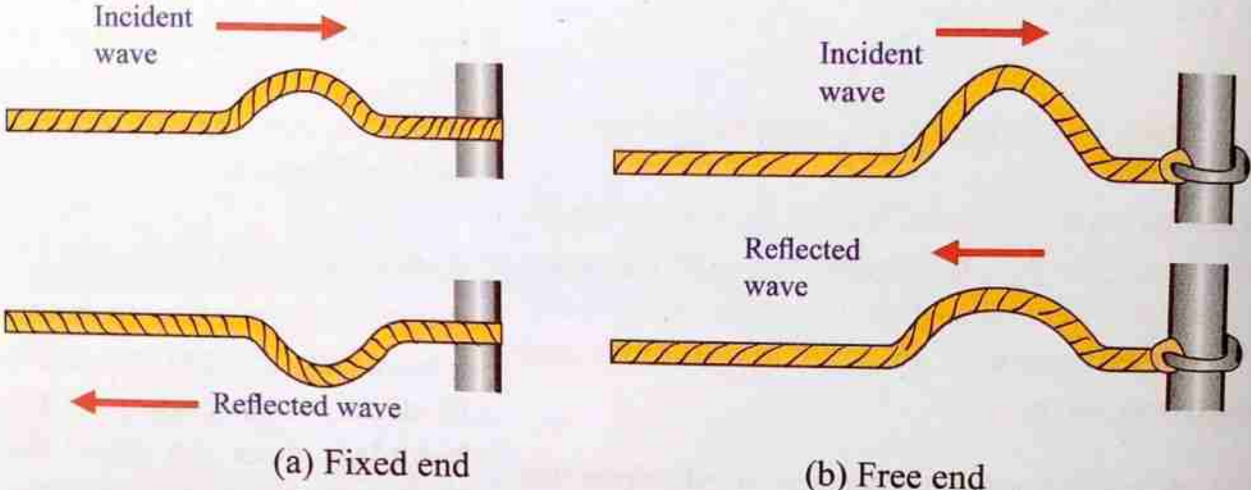


Figure 1.10: *Reflection of waves at fixed and free ends of a string*

When a wave encounters a boundary that allows it to pass through, part of the wave will be reflected and part will be transmitted into the new medium. Consider two ropes of different thicknesses tied together end-to-end and suppose that, a transverse wave is produced in the thinner rope. When the wave reaches the boundary between the two ropes, it will split into an inverted reflected wave and an upright transmitted wave. The reflected wave will have the same speed and wavelength as the incident wave. The transmitted wave will have a lower speed and a shorter wavelength than the incident wave. Each wave will have an amplitude less than that of the incident wave since the energy of the incident wave is split into the two waves. See Figure 1.11.

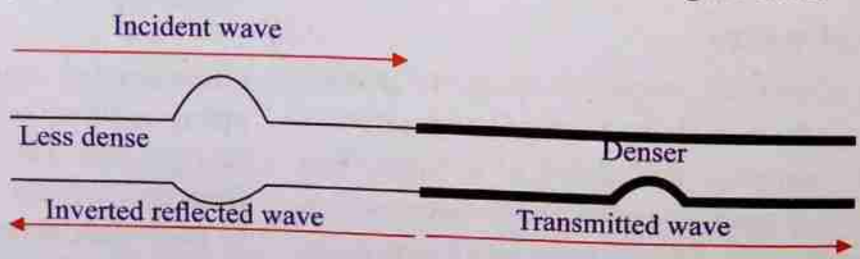


Figure 1.11: *Reflection of a wave travelling from a less dense medium to a denser medium*

If the new medium has a lower density, the reflected wave will not be inverted, as illustrated in Figure 1.12. It will have the same speed and wavelength as the incident wave. The transmitted wave will have a higher speed and longer wavelength. According to the principle of conservation of energy, when the wave breaks up into a reflected wave and a transmitted wave at the boundary, the sum of the energies of these two waves must be equal to the energy of the incident wave. Because the reflected wave contains only part of the energy of the incident wave, its amplitude must be smaller. Reflection and other behaviours of waves may be demonstrated using water waves. This is done using a ripple tank.

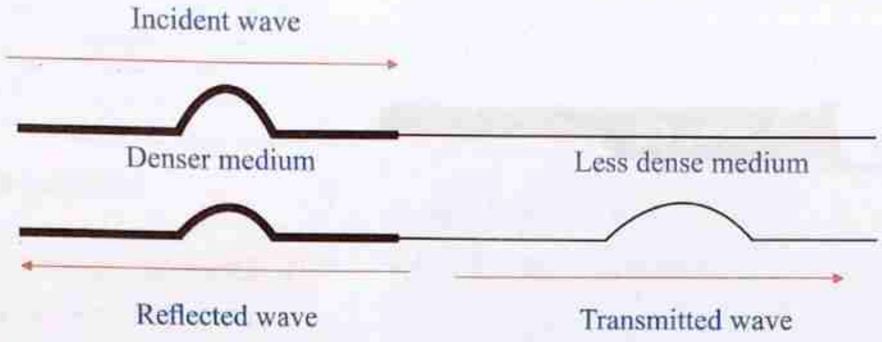


Figure 1.12: *Reflection of a wave travelling from a denser medium to a less dense medium*

Ripple tank

A ripple tank is an example of an instrument used to demonstrate the behaviour of waves. The structure of a ripple tank is shown in Figure 1.13. It consists of a power supply used to run an electric motor. When the motor runs it makes the oscillating paddle attached to an elastic band to vibrate on the water surface. The vibration of the paddle generates parallel water waves (ripples). The oscillating paddle is used to transform mechanical energy generated by the motor to ripples in a shallow tank of water. A bulb/lamp shines light through the water and a shadow of the wave pattern is produced on a sheet of paper or glass placed under the tank. The paper or glass act as viewing screen. All behaviours of waves can be demonstrated with the aid of a ripple tank.

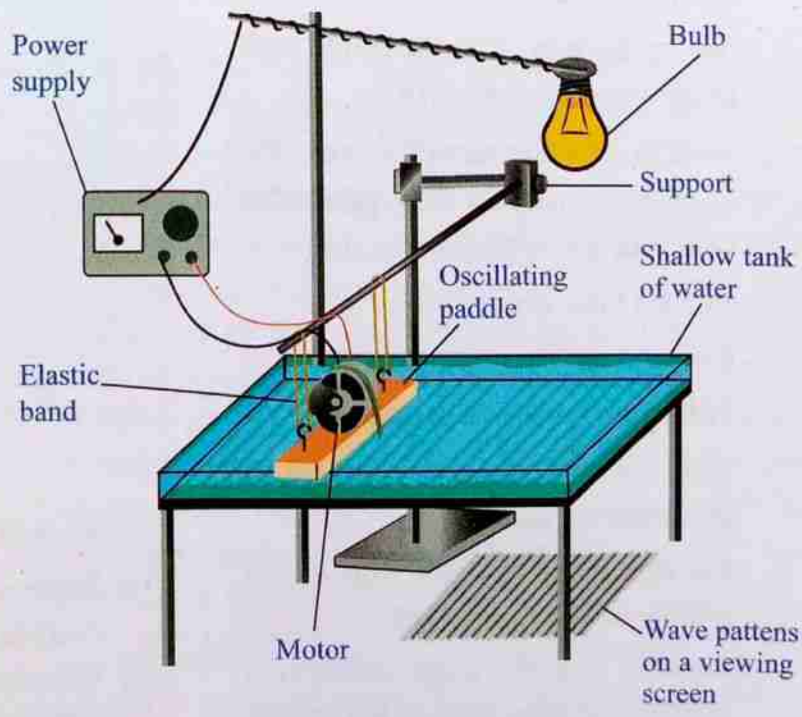


Figure 1.13: *Ripple tank*

Reflection of water waves can be observed by placing various obstacles in the tray of the ripple tank. The depth of the water can also be varied by laying glass plates of different thicknesses in the tray. This allows the observation of waves travelling from one medium to another.

Activity 1.1

Aim: To observe the reflection of water waves in a ripple tank.

Material: Ripple tank, barriers of different shapes (straight barrier, wooden block, concave and convex barriers), and droppers

Procedure

1. Assemble the ripple tank as illustrated in Figure 1.13.
2. Set up the barrier in the ripple tank as shown in Figure 1.14
3. Switch ON the motor to run the oscillating paddle that generates some waves in the ripple tank.
4. Record your observations.
5. Remove the oscillating paddle.
6. Put some water into a dropper.
7. Hold the dropper about 2 cm above the water surface.
8. Let one drop of water fall at the middle of the ripple tank.
9. Observe how the water waves will be reflected after striking the barrier.

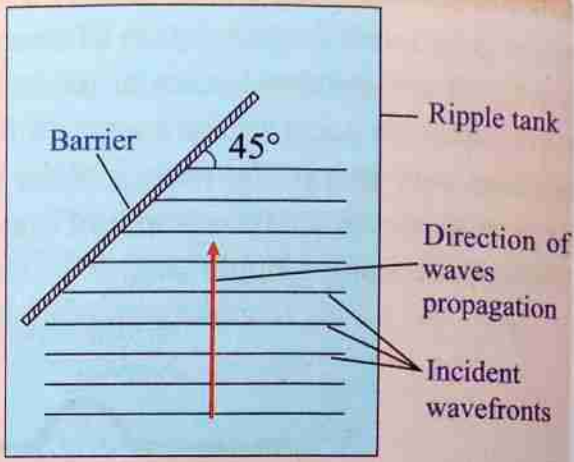


Figure 1.14

10. Draw the incident and the reflected wave patterns in steps 4 and 9.
11. Change the angle of the barrier and repeat steps 3-9. Observe any changes in the reflected waves.
12. Replace the rectangular barrier with a curved one (concave) as shown in Figure 1.15, and repeat steps 3-9.

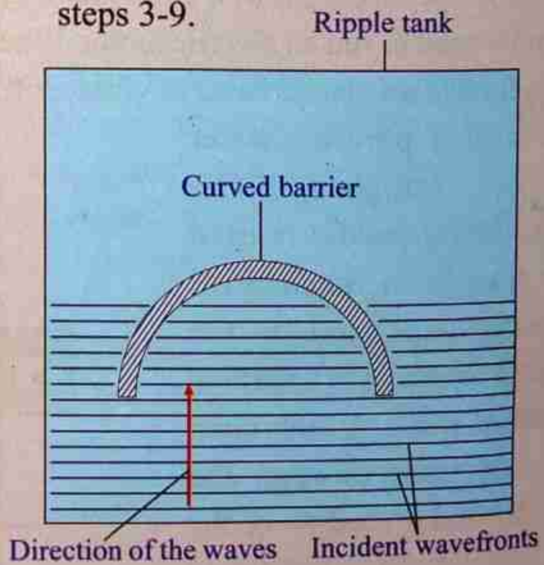


Figure 1.15

13. Record all your observations.
14. Replace the concave barrier with a convex one and repeat steps 3-9.
15. Observe any changes in the reflected waves.

Questions

(a) What happens to the incident water waves when they reach the rectangular barrier?

(b) Describe the wave pattern produced when the incident water waves strike the concave or convex barriers.

Reflection involves a change in the direction of waves when they fall on a barrier.

The direction in which a wave is travelling is represented by an arrow. The arrow is called a ray and is drawn perpendicular to the wavefronts. Upon reaching the barrier placed within the water, water waves bounce off the barrier and head in a different direction. Regardless of the angle at which the wavefronts approach the barrier, the waves will always be reflected in such a way that the angle of incidence at the barrier with respect to the normal is equal to the angle at which the waves are reflected off the barrier (Figure 1.16). This is in accordance to the laws of reflection which states that:

1. Upon reflection from a straight barrier, the angle of the reflected ray, r is equal to the angle of incident ray, i with respect to the normal, N (a line that is perpendicular to the surface at the point of contact). That is, $i = r$.
2. The incident line of propagation, the normal line and the reflected line of propagation all lie in the same plane.

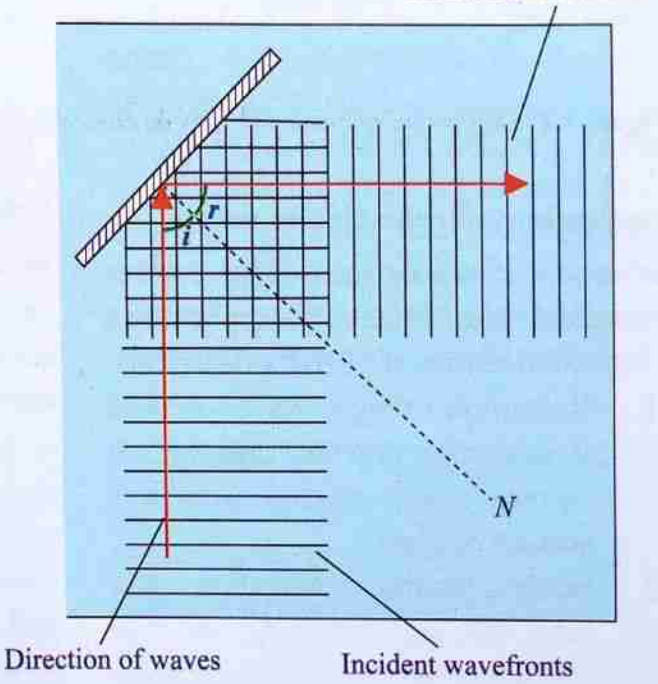


Figure 1.16: *Reflection of water waves*

When, a *straight water wave* strikes a curved barrier, the principles of reflection still apply, but the pattern becomes more complex. Consider a rubber tube having the shape of a parabola placed within the water. Upon reflection on the parabolic barrier, the water wave will change direction and head towards a point known as the focal point. This is the point at which the wave energy concentrates. After passing through the focal point, the waves spread out as shown in Figure 1.17. This is also the case when circular water waves strike a straight or an outward curved (convex) barrier.

Note that the parabolic barrier focusses the water waves exactly at half the distance from the centre of curvature.

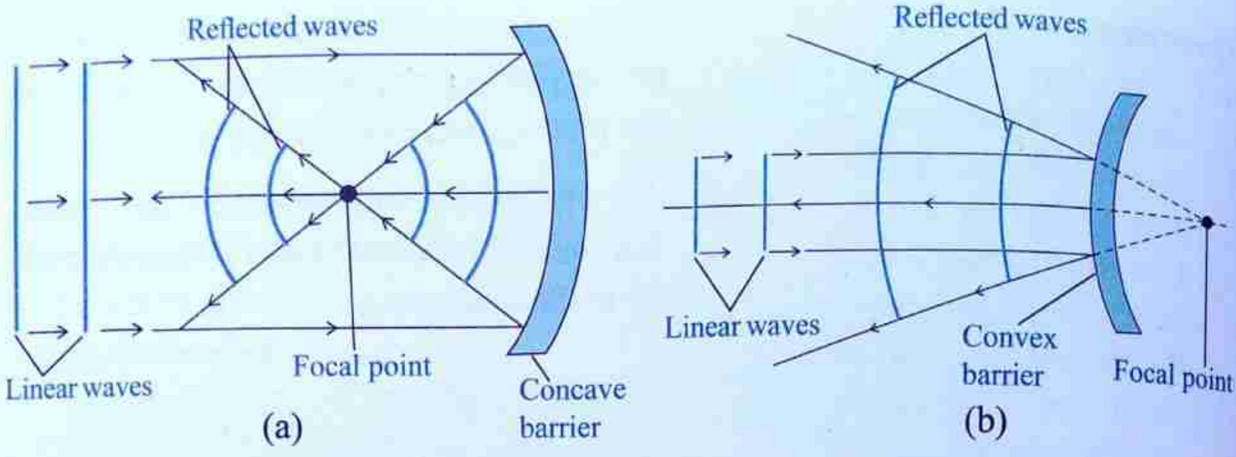


Figure 1.17: *Reflection of linear waves from curved barriers (a) concave barrier (b) convex barrier*

Applications of reflection of waves

Reflection of waves is used in various human activities. Some of the applications of wave reflections are hereby described.

1. Reflection of light waves is used in designing mirrors. Light waves bounce upon striking a silvery surface of glass.
2. Sonar (sound navigation and ranging) systems rely on the reflection of sound waves to measure the distance and speed of underwater objects.
3. The reflection of sound is what makes a hearing aid operate. Sound waves are reflected into a smaller region in a hearing aid, which directs the sound to the ear.
4. The soundboard is built upon sound reflection. Here, sound waves are uniformly reflected in an auditorium. This aids in the enhancement of sound quality.
5. The working of a stethoscope is based on the reflection of sound. In the stethoscope, the sound of a patient's heartbeat reaches a doctor's ear by multiple reflections of sound.

Refraction of waves

When waves travel from one medium to another, they tend to change their travelling speed. This phenomenon is known as refraction of waves. Refraction of waves occurs because the speed of a wave depends on the medium through which the wave is travelling. If the medium is changed, the speed of the wave will also change. The change in speed results in a change in the wavelength of the wave.

Activity 1.2

Aim: To observe the refraction of water waves in a ripple tank.

Materials: Ripple tank and its accessories, rectangular glass plates and metre rule

Procedure

1. Set a ripple tank as illustrated in Figure 1.13
2. Fill the ripple tank with water.
3. Gently place a rectangular glass plate in one part of the ripple tank to make it shallower than other parts

of the tank. The glass plate should be placed at the end opposite to that of the vibrator.

4. Produce some water waves using the vibrator.
5. Measure the distance between successive crests in the deeper part, λ and then in the shallow part, λ_1.

Questions

(a) Work out the ratio $\frac{\lambda}{\lambda_1}$.

(b) What is the relationship between the ratio of wave velocities in the deep and shallow parts of the tank?

When water waves travel from a deep part to a shallow part, the wavelength decreases as illustrated in Figure 1.18. However, the frequency of the waves does not change. Since velocity, v, is given by λf, the velocity of the waves decreases with decrease in wavelength.

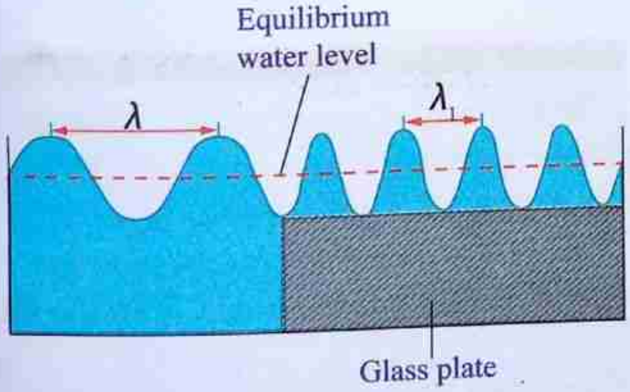


Figure 1.18: *Change in the wavelength of water waves*

The ratio $\frac{\lambda}{\lambda_1}$ is equal to the ratio of the velocities of water waves in the deep water, v to that in the shallow water, v_1, that is, $\frac{\lambda}{\lambda_1} = \frac{v}{v_1}$.

The velocity of water waves is higher in the deep water than in the shallow water.

Refractive index

The refraction of water waves can be further observed if the boundary between the deep and shallow regions is at a certain angle to the incident wavefronts. In this case the deep region acts as the first medium while the shallow region acts as the second medium. If in Activity 1.2, the glass plate was placed at some angle, water waves would be refracted as shown in Figure 1.19.

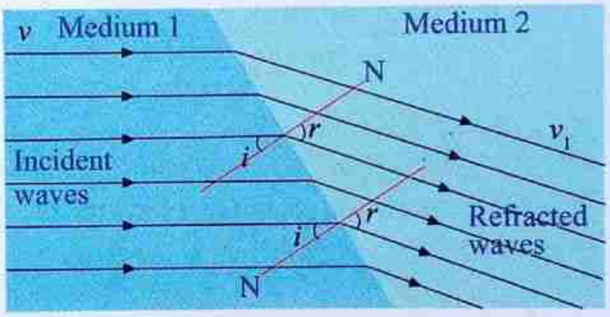


Figure 1.19: *Refraction of water waves*

Experimental observations have shown that water waves obey Snell's Law of refraction. That is,

$$\eta = \frac{\sin i}{\sin r}$$

where η is the refractive index of the second medium relative to the first medium, i is the angle of incidence and r is the angle of refraction.

The refractive index, η is also the ratio of the velocity of the wave in the first medium to that in the second medium, that is;

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

Example 1.4

The speed of light is $2.25 \times 10^8 \text{ m s}^{-1}$ in water and $3 \times 10^8 \text{ m s}^{-1}$ in air.

Determine:

(a) The refractive index of light from air to water.

(b) The angle of refraction in the water if the incident angle of light at the surface of water is 30°.

Solution

(a) Refractive index,

$$\eta = \frac{\text{Speed of light in air}}{\text{Speed of light in water}}$$

$$\eta = \frac{3 \times 10^8 \text{ m s}^{-1}}{2.25 \times 10^8 \text{ m s}^{-1}}$$

$$\eta = 1.33$$

Therefore, refractive index of water is 1.33.

(b) $$\eta = \frac{\sin i}{\sin r}$$

$$\sin r = \frac{\sin i}{\eta} \Rightarrow \frac{\sin 30°}{1.33}$$

$$r = \sin^{-1}(0.3760)$$

$$r = 22.08°$$

Therefore, the angle of refraction in the water is 22.08°.

Applications of refraction of waves

As in the case of wave reflection, wave refraction is also useful in many activities. The following are some applications of refraction of waves:

1. Refraction is used in optical instruments which focus or spread light. These include cameras, microscopes and telescopes.
2. Spectacles worn by people with visual impairment use the principle of refraction of light.
3. Since every material has its own value of refractive index, purity of a material can be identified by determining the refractive index of the material.

Interference of waves

Consider two or more waves propagating in the same medium. If the waves are travelling in the same direction, they tend to add up forming a new wave of larger amplitude. If waves of the same frequency and amplitude travelling with the same speed but in opposite directions meet, they tend to cancel each other. That is, no new wave is formed. The addition of waves is called wave superposition and is governed by the principle of superposition of waves.

The principle of superposition of waves states that, the resultant disturbances at any point is equal to the algebraic sum of the disturbances of individual waves at that point.

When two or more waves combine to form a resultant wave, the waves are said to have interfered. Therefore, wave interference is a combination of two or more waves to form a resultant wave in which the particle displacement is either reinforced or cancelled. When two waves interfere and if a crest of one wave meets the crest of another wave at the same point, then the resultant wave has a larger amplitude, as shown in Figure 1.20. This is referred to as *constructive interference*. Constructive interference occurs at any location along the medium where two or more interfering waves have displacements in the same direction (the waves are in phase). On the other hand, if a crest of a wave meets a trough of another wave, then the resultant wave has a smaller amplitude, as shown in Figure 1.21. This is referred to as *destructive interference*. Destructive interference occurs at any point along the medium where the interfering waves have displacements in opposite directions (the waves are out of phase).

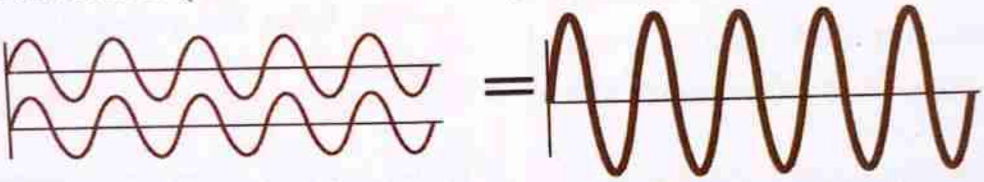


Figure 1.20: *Constructive interference*

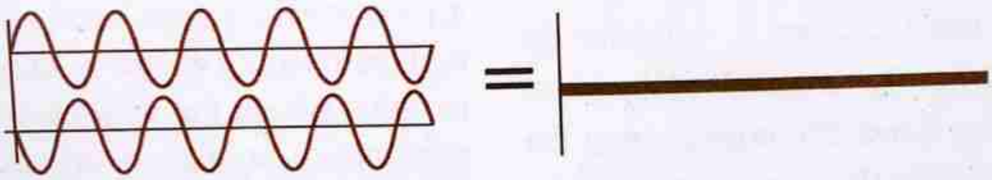


Figure 1.21: *Destructive interference*

Examples of the resultant displacement due to the interference of two waves of different amplitudes are given in Table 1.1.

Table 1.1 *Resultant displacements from interference*

Amplitude of wave A	Amplitude of wave B	Resultant amplitude
+1	+1	+2
+2	-1	+1
-1	+1	0
+1	-2	-1
-1	-1	-2

In constructive interference, we get lines of increased disturbance. These lines are called antinodal lines. In destructive interference, we get lines of zero disturbance. These lines are called nodal lines. Figure 1.22 shows circular waves from two dippers that are close together in a ripple tank. The waves cross through one another and result in both constructive and destructive interference.

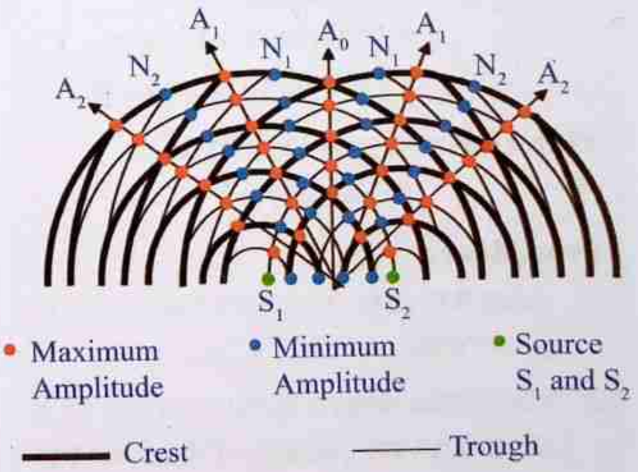


Figure 1.22: *Interference of circular waves*

The red dots indicate where crests or troughs meet giving rise to maximum amplitude which results to maximum intensity. The blue dots indicate where crest meets a trough giving rise to minimum or zero amplitude resulting to minimum intensity. The lines of maximum constructive interference are labelled A_0, A_1 and A_2. Points on these lines move up and down with higher amplitudes much than they would if the waves came from one source alone. The lines labelled N_1 and N_2 represent bands along which there is maximum destructive interference. Points on these lines move up and down with much less amplitude than they would if the waves came from either source alone.

Note that, if the two waves from S_1 and S_2 have the same amplitude, then the points along the lines A_0, A_1 and A_2 will have the amplitude twice the amplitude of the wave. On the other hand, the points along the lines N_1 and N_2 will have zero amplitude.

Activity 1.3

Aim: To investigate the interference of water waves in a ripple tank.

Materials: Ripple tank and its accessories

Procedure:

1. Attach two-point sources to the wave generator.
2. Generate some water waves at a constant frequency. Observe the pattern created by the waves.
3. Increase the frequency of the wave generator and observe what happens to the pattern created by the waves.
4. Decrease the frequency and again observe what happens to the pattern of the waves.
5. While keeping the frequency constant, increase the distance between the two-point sources.
6. Record all your observations.

Question

Describe the pattern produced by the waves from the two-point sources in each case.

Water waves from two identical point sources add up at certain points (where a crest meets a crest) and cancel out at certain points (where a crest meets a trough). Where the water waves add up, constructive interference occurs. The water waves have increased amplitudes along lines of constructive interference. On the other hand, destructive interference occurs where the water waves cancel out. Along the lines of destructive interference, the water is observed to be still.

Applications of interference of waves

1. Wave interference is applied when creating holograms. A hologram is a photograph of an interference pattern which is able to produce a three-dimensional image.
2. Destructive interference is used in noise-reduction systems such as earphones and car mufflers. The systems capture sound from the environment and produces a second wave, which interferes with the first wave destructively leading to the reduction in the loudness of the noise.
3. Another application of interference of waves is in Active Noise Control. This is based on the fact that the wave generated by a primary source (such as an engine) can be cancelled by the

wave emitted by a secondary source (loudspeakers) driven at the same frequency as the primary source, so that the two waves cancel out each other. This technique is applied to reduce the annoying propeller noise inside the cabin of an aircraft.

Diffraction

Suppose water waves in a ripple tank are allowed to pass through a gap (aperture) formed by an obstacle (Figure 1.23). If the width of the gap is wider than the wavelength of the water waves, the waves appear to move in the same straight line as the incident wave. If the width of the gap is decreased to the order of the wavelength of water waves, then after the wave has passed through, it spreads in all directions. The spreading of the wave as it encounters an obstacle is called *diffraction*. The extent of spreading depends on the width of the aperture in comparison to the wavelength of the incident wave. A waves whose wavelength is comparable to or larger than the width of an aperture spreads out in all forward directions upon passing through the aperture. The bigger the width, the less the diffraction.

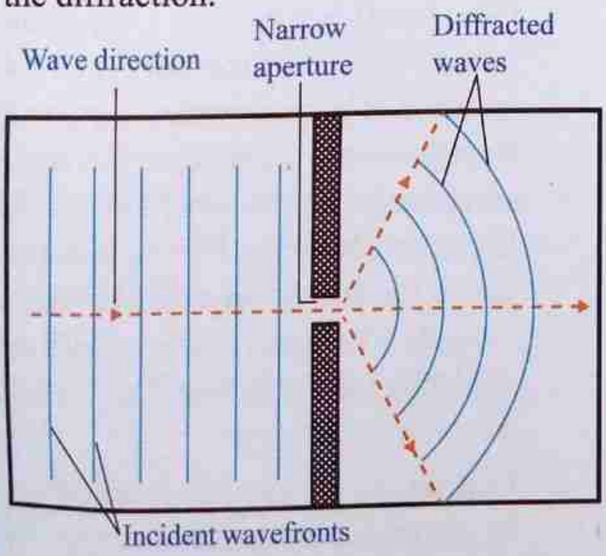


Figure 1.23: *Diffraction of water waves*

Activity 1.4

Aim: To investigate diffraction of water waves in a ripple tank.

Materials: Ripple tank and its accessories, glass block and pencil

Procedure

1. Add water to the ripple tank to a depth of approximately 1 cm.
2. Adjust the legs of the ripple tank to make the depth of the water as uniform as possible.
3. Place a glass block in the water across the tank.
4. Drop a pencil onto the water, just a few centimetres behind the glass block.

   Note: The pencil should be aligned in such that the generated waves travel in a direction perpendicular to the edge of the glass block.
5. Observe the waves as they move past the glass block.
6. Put some additional glass blocks in the ripple tank to create a barrier across the entire tank. Leave an opening of about 1 cm at the centre of the ripple tank.
7. Generate some straight waves using the oscillating paddle of the ripple tank.
8. Observe what happens to the waves as they travel past the barrier.
9. Increase the size of the opening, then repeat step 6.
10. Replace the straight wave source with a point source, then repeat step 6.
11. Record all your observations.

Questions

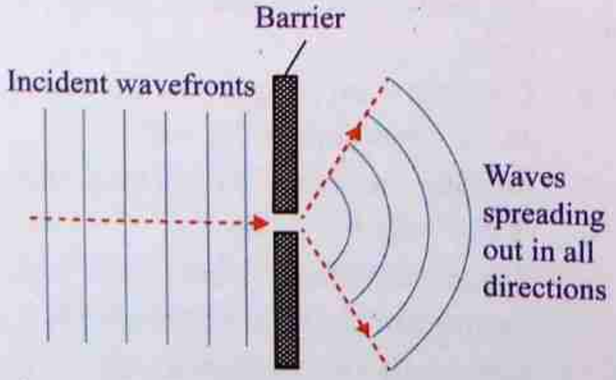
(a) Explain what happens to the waves on passing through the single block barrier.

(b) Explain what happens to the waves on passing through the narrow opening.

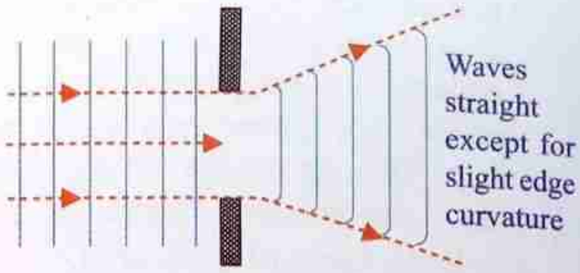
(c) What was the effect of increasing the size of the opening on the waves passing through it?

When straight waves are approaching a barrier, the barrier obstructs a part of the wave that strikes it and allows just that part of the wave which does not strike it to pass through. The wavefronts that pass the barrier spread into the shadow area of the barrier. The spreading takes place whenever a wave meets a barrier or an aperture (narrow opening).

When the gap is narrow, the straight wavefronts are converted into circular wavefronts as illustrated in Figure 1.24 (a). These wavefronts appear to be produced by a new point source in the gap. They spread out round the edges of the opening in all directions. When the gap is wide as in Figure 1.24 (b), the waves emerge almost straight, apart from a slight curvature, and spread out at the edges.



(a) Narrow gap



(b) Wide gap

Figure 1.24: *Diffraction of water waves*

The amount of diffraction is maximum when the width of the opening is the same as the wavelength of the waves. Diffraction can occur with any kind of wave. For example, ocean waves diffract around obstacles just as sound and light can diffract around objects. The amount of diffraction depends on the size of the aperture and the wavelength of the incident wave. Since sound has longer wavelength than light, it diffracts more than light upon falling on an aperture.

Applications of diffraction of waves

Diffraction of waves has many applications. Some of these applications are:

1. The process of diffraction is significantly used in long-distance radio signal propagation. Despite the curved surface of the Earth and the presence of huge obstacles on it, radio signals from a transmitter can reach an observer on the other side of the Earth or obstacle. This is possible due to the diffraction of waves at the obstacle. The signal diffracts to fill the void after the obstacle and surfaces to travel to the observer.
2. Diffraction is used in a hologram to generate a three-dimensional impression of an image. Different

versions of the image get diffracted and reach a lens from multiple sides, all together forming an interference pattern. This pattern is then made to fall on a holographic plate providing a three-dimensional image.

3. Diffraction is used in measuring the coefficient of thermal expansion, crystallite size and thickness of thin films.
4. X-rays diffraction is used to determine the distance between two consecutive atoms of a material. X-ray diffraction process is crucial in the meteorological, pharmaceutical, chemical, and other related industries because whenever researchers come across some unidentified materials, they need to figure out the details about their structure, starting with the alignment, distance, and other characteristics of their atoms.

Exercise 1.1

1. If you want to hit a fish under water using a spear, should you aim below it, straight at it or above it? Explain.
2. A periodic disturbance in a lake creates waves which emanate outwards from its source to produce circular wave patterns. If the frequency of the source is 2 Hz and the wave speed is 5 m s^{-1}, determine the distance between adjacent wave crests.
3. A wave whose speed in the first medium is 4.4 m s^{-1} enters a second medium. The wavelength changes from 2 m to 3 m. What is the speed of the wave in the second medium?
4. For a certain transverse wave, the distance between two successive crests is 1.2 m, and eight crests pass a given point along the direction of travel every 12 s. Calculate the wave speed.
5. A sinusoidal wave is travelling along a rope. The oscillator that generates the wave completes 40 vibrations in 30 s. Also, a given crest of the wave travels 425 cm along the rope in 10 s. Calculate the wavelength of the sinusoidal wave.
6. The velocity of light in water is $2.2\times10^{8}\,\text{m}\,\text{s}^{-1}$ and the velocity of light in glass is $2.0\times10^{8}\,\text{m}\,\text{s}^{-1}$.

Calculate:

(a) The relative refractive index as light passes from water to glass.

(b) The angle of incidence in the water which would produce an angle of refraction of 30° in the glass.

7. A straight vibrator causes water ripples to travel across the surface of a shallow tank. The ripples travel a distance of 33 cm in 1.5 s and the distance between successive crests of the wave is 4.0 cm. Calculate the frequency of the vibrator.
8. Radio and light waves travel at a velocity of $3\times10^{8}\,\text{m}\,\text{s}^{-1}$ in air. Calculate:

   (a) The wavelength of radio waves when transmitted at a frequency 150 MHz; and

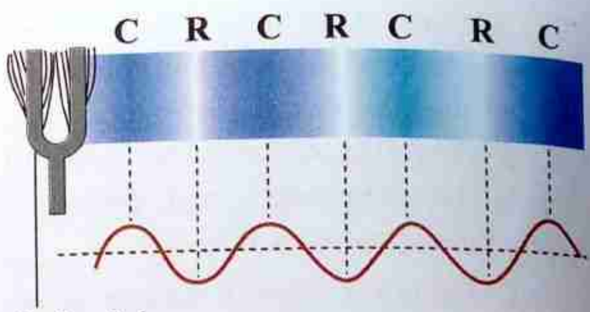
(b) The velocity of light in glass of refractive index 1.5.

9. Water ripples are caused to travel across the surface of a shallow tank by means of a suitable straight vibrator. The distance between successive crests and troughs is 1.5 cm and the wave travels 25.2 cm in 1.2 s. Calculate the wavelength and the velocity of the waves and the frequency of the vibrator.

10. If a wave has a velocity of 330 m s^{-1} and a wavelength 0.5 m, calculate the frequency of the vibrator producing the wave.

Sound waves

When the air is set to vibrate by an oscillating body such as a tuning fork, string, whistle, or clarinet, it produces sound. Sound is a mechanical vibration transmitted through a medium such as solids, liquids or gases. Unlike waves in a string which move in one dimension and water waves which move in two-dimension, sound waves move in three dimensions. It is a longitudinal wave as illustrated in Figure 1.25. As the tines of the tuning fork vibrate, they set the surrounding air molecules into vibration. As neighbouring molecules interact, the vibrations travel away from the tuning fork in all directions. Because of the longitudinal vibrations of the air molecules, there are regions where the molecules are compressed, and adjacent regions where they are spread out. Thus, sound waves are sometimes referred to as pressure waves.



C = Compression **R** = Rarefaction

Figure 1.25: *Sound waves generated by vibrating tuning fork*

Propagation of sound waves

Like all mechanical waves, sound waves require a medium to be transmitted. Sound is transmitted by vibration of particles of the medium. One particle vibrates to transfer energy to the next until the sound reaches another point. Sound travels quicker when the particles are closer together. For example, sound travels faster in solids than in gases. This is because molecules of a solid are more closely packed than gas molecules. If one stands near a railway line and another person taps the rail some distance away, two successive sounds will be heard, the first through the rail and a later one through the air.

Sources of sound waves

Sound waves are produced by almost everything including people, animals, plants and machines. Musical instruments are designed to produce specific types of sound. These instruments include guitars, violins, pianos, organs, flutes, drums and xylophones. Figure 1.26 shows some musical instruments.

Figure 1.26: *Musical instruments*

Categories of sound waves

The human ear is very sensitive and can detect even faint sounds. Whether you hear or do not hear a sound depends on its loudness and frequency. The average human ear can detect sounds in the frequency range of 20 Hz to 20 000 Hz. Nevertheless, due to various reasons including age, there are significant differences between individuals, especially at high frequencies. The upper limit of the audible range decreases throughout an adult's life. The range of sound frequency which can be detected by the human ear is known as the *audible range* or *audio range*. Waves which lie in this range are called audible waves.

The ear is most sensitive to sounds with a frequency around 3 000 Hz. Sound waves with frequencies below 20 Hz are said to be *infrasonic*. Conversely, sound waves that are above 20 000 Hz are called *ultrasonic*. Elephants communicate using infrasonic waves. Some animals including bats and dolphins can detect ultrasonic sounds with frequencies as high as 100 000 Hz. Ultrasonic waves are also applied in medical imaging devices such as ultrasound machines. The average human ear can distinguish between two simultaneous sounds if their frequencies differ by at least 7 Hz. Table 1.2 shows the audio range for different animals.

Table 1.2: *Audio range for different animal species*

Animal	Audibility range (Hz)
Bat	9 000 – 200 000
Blue whale	10 – 40
Cat	45 – 64 000
Cattle	23 – 35 000
Chicken	10 – 12 000
Dog	67 – 45 000
Elephant	14 – 16
Horse	55 – 33 500
Owl	200 – 12 000
Penguin	100 – 15 000
Rabbit	360 – 42 000
Rat	200 – 76 000
Risso's dolphin	8 000 – 100 000
Sheep	100 – 30 000

The human ear

The human ear converts sound energy to mechanical energy and then to electrical energy which acts as a signal sent to the brain via nerves. Human ears can discriminate between sound based on frequency, amplitude and direction. The human ear consists of three basic parts: the outer ear, the middle ear and the inner ear as shown in Figure 1.27.

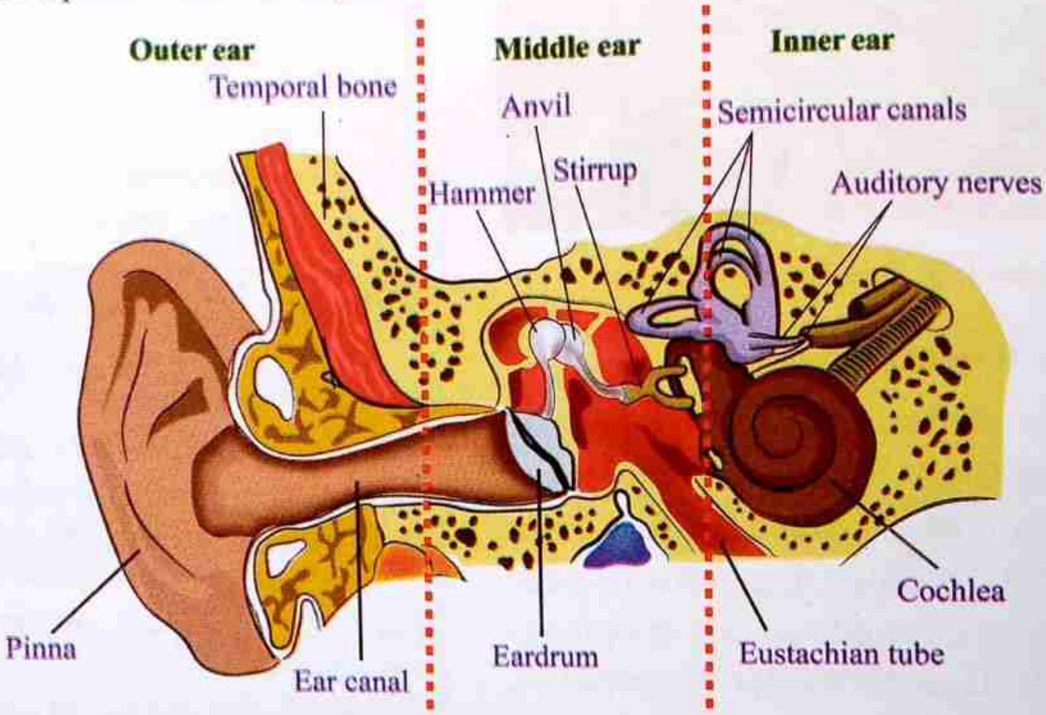


Figure 1.27: *The human ear*

The outer ear

The outer ear consists of the pinna and the ear canal. The outer ear channels sound waves through the *ear canal* to the *eardrum* of the middle ear. In the outer ear, the sound is still in the form of a pressure wave, with an alternating pattern of high- and low-pressure regions.

The middle ear

The middle ear is an air-filled cavity that consists of an eardrum and three small interconnected bones, the *hammer, anvil* and *stirrup*. The *Eustachian tube* connects the middle ear to the throat. Its purpose is to regulate pressure.

A compression of the incoming sound wave forces the eardrum inward and a rarefaction allows the eardrum to move outward. In this way, the eardrum vibrates at the same frequency as an incoming sound wave. The movements of the eardrum set the hammer, anvil and stirrup into motion. The three tiny bones amplify the vibrations of the incoming sound wave. Because the stirrup is connected to the inner ear, the vibrations are transmitted to the fluid of the inner ear.

The inner ear

The inner ear consists of the cochlea, the semicircular canals and the auditory nerve. The cochlea and the semicircular canals are filled with a water-like fluid. The fluid and nerve cells of the semicircular canals help in maintaining the body balance. The inner surface of the cochlea is lined with hair-like nerve cells that differ in length. The nerve cells have different degrees of resilience to the fluid which passes over them.

As a compressional wave moves from the interface between the hammer of the middle ear and the oval window of the inner ear through the cochlea, the small hair-like nerve cells are set in motion. Each hair cell has a natural sensitivity to a particular frequency of vibration. When the frequency of the incoming wave matches the natural frequency of the nerve cell, that nerve cell vibrates with larger amplitude. The increased vibration makes the cell release an electrical impulse which passes along the auditory nerve to the brain for interpretation.

Echo and reverberation

Sound, like any other waves, can be reflected off a flat or hard surface and obeys the same laws of reflection. The reflection of a sound leads to the formation of an echo.

Echo

Suppose you are standing about 120 metres from a high building. If you clap your hands, after sometime you will hear again the sound of clapped hands. The repetitive sound that you hear is called an echo that reaches the ear more than 0.1 s after the original sound was heard. At this time interval, the sensation of the original sound will have died out and the reflected sound will be heard as a distinct sound. The memory of a sound persists in the brain for approximately 0.1 s. This means that the original sound and the reflected sound must be separated by a time interval of 0.1 s for the echo to be interpreted by the brain as a distinct sound.

In this phenomenon, the sound must first travel to the obstacle. When the sound returns, it covers the same distance. The minimum time that the sound should take to reach the obstacle for an echo to occur is $\frac{0.1\text{ s}}{2} = 0.05\text{ s}$. Figure 1.28 illustrates the sound wave travelling to an obstacle and then reflected back to the observer's ear.

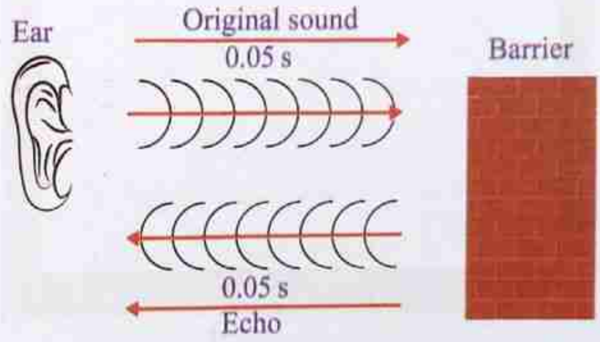


Figure 1.28: *Formation of an echo*

At room temperature, the speed of sound in air is about 340 m s^{-1}. Since distance is given by $distance = speed \times time$, the minimum distance, d to the obstacle for an echo to occur is given by:

$$\begin{aligned} d &= speed \times time \\ &= 340\text{ m s}^{-1} \times 0.05\text{ s} \\ &= 17\text{ m} \end{aligned}$$

An obstacle must be at least 17 m away for a distinct echo to be heard. However, materials through which the sound propagates may affect the occurrence of an echo.

Reverberations

When a sound is produced in an enclosed space with dimensions that are approximately less than 17 m, multiple reflections occur, forming multiple echoes picked by the ear. Since the dimensions of the enclosed space are less than 17 m echoes can reach the ear in less than 0.1 s. These echoes reach the ear of the listener and superpose to produce a louder and more sustained sound. This phenomenon is referred to as reverberation and is illustrated by Figure 1.29.

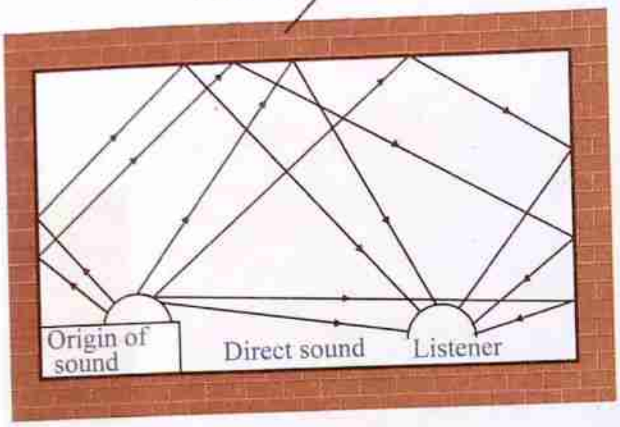


Figure 1.29: *Reverberations*

Speed of sound in air

The speed of sound in air can be determined using echo. If sound travels to a wall and back, the total distance covered by the sound in time t is $2d$.

Therefore, the speed of sound waves in air is given by: $v = \frac{2d}{t}$ m s^{-1}

Suppose the distance, d to the wall is 500 m and the time taken, t for the echo to be heard is 3 s. The speed of sound in air will be:

$$Speed = \frac{2 \times distance}{time}$$

$$= \frac{500\text{ m} \times 2}{3\text{ s}}$$

$$= 333.3\text{ m s}^{-1}.$$

It is known that, the speed of sound in dry air is calculated using the relationship:

$$v = 331.4 + 0.6\,\theta\text{ m s}^{-1}$$

where θ is the air temperature in centigrade at which the speed is measured.

For example, the speed of sound in air at 25 °C is given by:

$$v = 331.4 + 0.6\,\theta$$

$$= 331.4 + (0.6 \times 25)$$

$$= 331.4 + 15$$

$$= 346.4\text{ m s}^{-1}.$$

Activity 1.5

Aim: To determine the speed of sound in air by echo method.

Materials: Two wooden blocks, stopwatch, measuring tape, vertical cliff or wall

Procedure:

1. Check your school dining hall so that it is not occupied by other students.
2. Close all windows and doors of the building.
3. Mark one end of the building as the starting point.
4. While in the building, move close to the starting point and hit the two wooden blocks together. Note if an echo occurs.
5. Change your position and repeat step 4.
6. Repeat step 5 until you hear an echo.
7. Mark the point where the echo occurs as shown in Figure 1.30.
8. Measure the distance from point A to the wall.

9. While standing at point A let one of you hit the block and the other measure the interval from the time of hitting the block to the time an echo is heard.

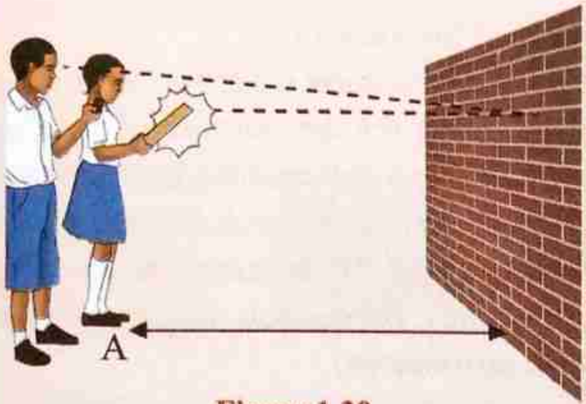


Figure 1.30

Question

Determine the speed of sound, v in air.

Exercise 1.2

1. Sound of explosions taking place on other planets is not heard by a person on the Earth. Explain.
2. A sonar device on a submarine sends out a signal and receives an echo 5 s later. Calculate the speed of the sound in water if the distance of the object from the submarine is 3 625 m.
3. A person standing 99 m from the foot of a tall cliff, claps hands and hears an echo 0.6 s later. Calculate the speed of sound in air.
4. An observer stands between two distant cliffs and claps hands. An echo is received after 2 s and 2.5 s respectively. If the speed of sound in air is 330 m s^{-1}, find the distance between the cliffs.
5. How long would it take for a 30 Hz beat to reach an audience member 100 m away when the ambient temperature is 21 °C?
6. Calculate the ratio of velocities of sound produced in dry air at 42 °C to another at 60 °C.
7. Two persons stand facing each other, 200 m apart on one side of a high wall and at the same perpendicular distance to it. When one fires a pistol the other hears a report 0.60 s after the flash and a second 0.25 s after the first. Explain this and calculate:
   (a) The velocity of sound in air.
   (b) The perpendicular distance of the persons from the wall.
   (c) Draw a diagram showing the positions of the persons and the wall.

Musical sounds

When listening to music from a radio, one can differentiate the sounds coming from saxophones, drums, guitars, violins and other musical instruments. This is possible because the patterns of the sound waves from each source is different.

That is, sounds from different instruments can be classified using the properties of musical sounds.

Properties of musical sounds

The musical sounds produced by different musical instruments have distinct properties that are used to describe them. These properties include *loudness*, *pitch* and *timbre*.

Loudness

When listening to music, there is a possibility of sensing high sound and low sound. This happens depending on the energy of the sound that enters the ears per second.

If the energy is increased the music becomes louder. If the sound energy is decreased the music becomes less loud. The loud sound is obtained when the amplitude of a vibrating source is high. Thus, the higher the amplitude the louder the sound and vice versa.

Pitch

The frequency of a sound wave approaching the ear determines how low or high the sound is. The scale of high or low for sound is termed as pitch. The sound itself is also called a note. The pitch is at high scale if the frequency is high. It is also called a high pitch note.

Timbre

If the same note sounds differently on different instruments, it is said that the two instruments have different timbre or quality. The difference is due to the fact that, with the exception of tuning forks and function generators, other instruments cannot emit pure, musical sound at the same frequency. The sound (note) consists of a fundamental frequency mixed with other frequencies called overtones. Overtones have frequencies that are exact multiples of the fundamental frequency. The quality or timbre of sound is determined by the strength of overtone.

Musical instruments

A musical instrument is a device constructed or modified for the purpose of making music. Musical instruments are grouped into three categories based on how they initially produce sound. These categories are *string instruments*, *percussion instruments* and *wind instruments*.

String instruments include the violin, piano and guitar. They produce sound from stretched strings that are plucked (guitar), or bowed (violin) or struck (piano). Figure 1.31 shows some of the string instruments.

(a) Guitar

(b) Violin

(c) Piano

Figure 1.31: *String instruments*

Percussion instruments produce musical sounds by being struck, shaken, rubbed, scrapped, or by any other action which sets the object into vibration. Musical instruments in this group include the drum, cymbals, tambourine and xylophone. Figure 1.32 shows some of the percussion instruments.

(a) Drum (b) Xylophone

Figure 1.32: *Percussion instruments*

Wind instruments are made up of a tube in which a column of air is set into vibration by the player blowing into (or over) a mouthpiece at the end of the tube. They include recorders, flutes, tuba and trumpets. See Figure 1.33.

(a) Tuba (b) Trumpets

(s) Flute

Figure 1.33: *Wind instruments*

Task 1.3

Construct a simple musical instrument of your choice. Classify the instrument into its respective category.

Stationary waves

Consider two transverse waves A and B with equal amplitudes, wavelengths and speeds travelling in opposite directions through a string. As the waves pass each other, their amplitudes combine. If a crest of wave A combines with a trough of wave B as Figure 1.34 illustrates, the two amplitudes cancel out and the net displacement of particles in the medium will be zero.

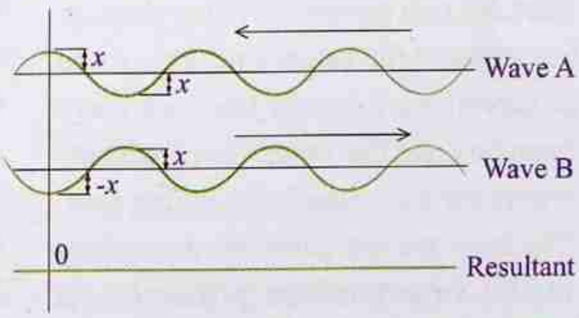


Figure 1.34: *Crest of wave A combines with a trough of wave B*

If a crest of wave A combines with a crest of wave B as in Figure 1.35, the result is a displacement which is twice larger than either A or B.

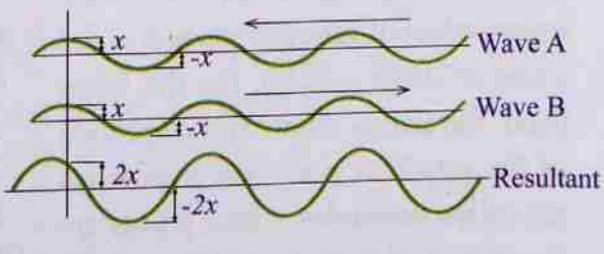


Figure 1.35: *Crest of wave A combines with a crest of wave B*

As the two waves pass along the string, their amplitudes will alternate between adding together and cancelling. The result is a wave in which particles oscillates back and forth but the wave does not propagate. Such a wave is called a *stationary* (or standing) wave.

Stationary waves can be produced in a string by means of vibrating tuning fork. If one end of the string is fixed to a stationary boundary and the other end to the tuning fork the string can move up and down due to the vibration of the tuning fork. The produced travelling waves reflect from the ends and travel in both directions of the string. Thus, there are two waves travelling through the string at the same time. One wave is travelling towards the stationary boundary and the other wave (reflected wave) travels towards the tuning fork. The two waves combine according to the superposition principle. At certain points along the string the two waves cancel out resulting to zero displacement, and at other points the waves add up to produce a maximum displacement. The points where the amplitude of the resultant wave is zero are called nodes. Nodes are independent of time (no motion in the string at these points). On the other hand, the points where the amplitude of the resultant wave are maximum are called antinodes. These points are dependent of time. A standing wave is illustrated in Figure 1.36.

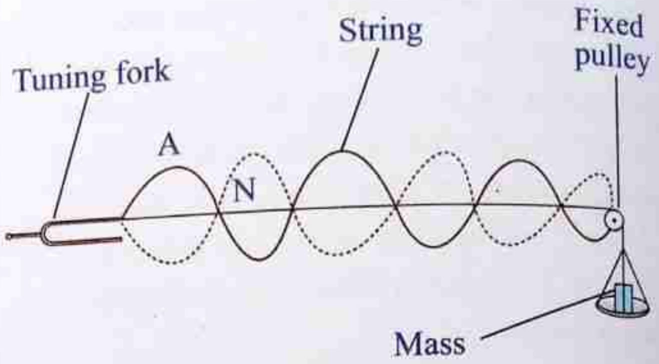


Figure 1.36: *Stationary wave*

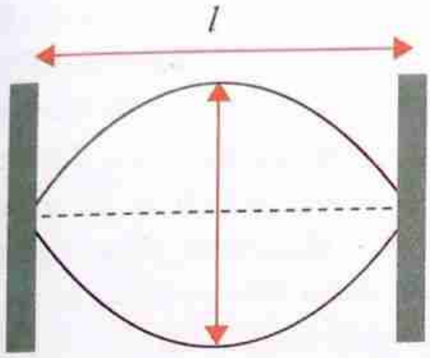
Note that, the profile of a stationary wave does not travel. That is, energy is not transmitted with the wave although there is energy associated with it.

Fundamental note, harmonics and overtones

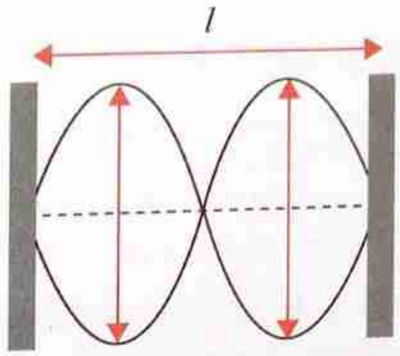
Stationary waves on a string that is fixed at both ends are restricted to having only certain wavelengths. The wave must "fit" between the ends of the string with a node at each end.

The lowest frequency that a vibrating string or pipe can produce is called the *fundamental frequency* and the corresponding note is called the *fundamental note*. A note whose frequency is *n* times (where *n* is a whole number) that of the fundamental note is called the n^{th} harmonic. The first harmonic is therefore the fundamental note. The fundamental note occurs when there are nodes at each end with a single antinode between them. The fundamental note is called the first harmonic. A similar behaviour can be observed on waves in pipes or tubes.

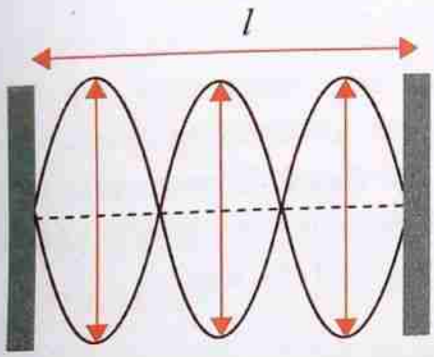
The overtones of a note are notes of a higher frequency which are produced with the fundamental note. The first overtone is the harmonic whose frequency is lowest among those with fundamental note. The second overtone is the next higher harmonic which is present. Higher harmonics occur when there are additional nodes as illustrated in Figure 1.37.



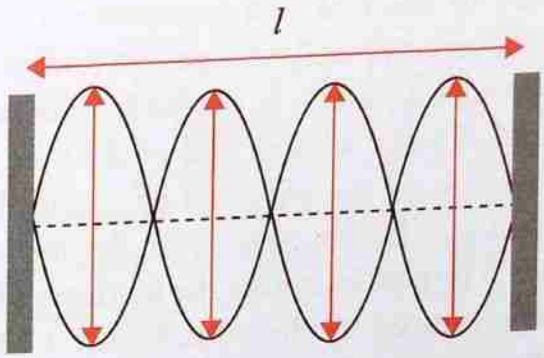
(a) Fundamental note, first harmonic



(b) Second harmonic (First overtone)



(c) Third harmonic (Second overtone)



(d) Fourth harmonic (Third overtone)

Figure 1.37: *Fundamental notes and harmonics*

The fundamental note in Figure 1.37 (a) consists of one half of a cycle. If the length of the string is l, then the wavelength λ_1 of the fundamental note is given by:

$$\frac{\lambda_1}{2} = l \Rightarrow \lambda_1 = 2l$$

The second harmonic in Figure 1.37 (b) has consists of one full cycle and has a wavelength, $\lambda_2 = l$.

The third harmonic in Figure 1.37 (c) is one and one-half cycles giving, $l = \frac{3}{2}\lambda_3$.

The wavelength will be given by, $\lambda_3 = \frac{2l}{3}$.

The fourth harmonic in Figure 1.37 (d) consists of two full cycle and has a wavelength, $l = 2\lambda_4$ giving $\lambda_4 = \frac{l}{2}$.

In general, for a string of length, l fixed at both ends, the wavelength, l_n of the n^{th} harmonic is given by: $\lambda_n = \frac{2l}{n}$.

Since $v = \lambda f$, the frequency of the n^{th} harmonic (f_n) is given by:

$$f_n = \frac{v}{\lambda_n}$$

Therefore, $f_n = \frac{nv}{2l}$

Note: The overtone is always one step ahead of harmonic. Thus, for n^{th} harmonic, the overtone is $(n+1)^{th}$.

Therefore, for harmonic, $f_n = nf_0$ while for overtone, $f_n = (n+1)f_0$, where $n = 1$, 2, 3,.....

Sonometer

A sonometer is a device that is used to study the properties of waves, particularly stationary waves. A wire or string is attached to one end of a sonometer. The opposite end is passed through two bridges and through a pulley. A weight hanger is suspended from the free end of the wire. The tension of the wire can be changed by placing different masses on the mass hanger. The box increases the loudness of sound produced by the wire. Figure 1.38 shows a sketch of a sonometer.

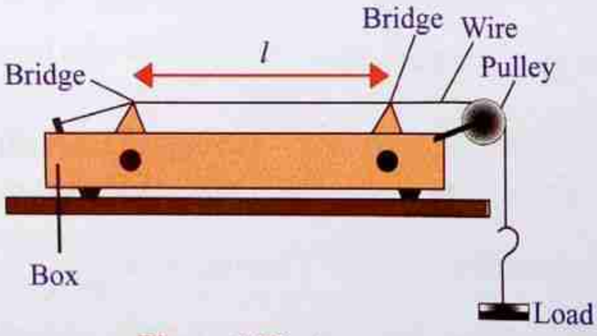


Figure 1.38: *A sonometer*

If the mid-point of the wire is plucked, the middle will form an antinode while the two fixed ends will have nodes as indicated in Figure 1.39.

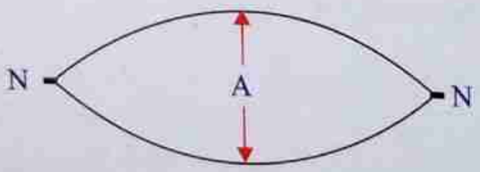


Figure 1.39: *Node and antinode for a sonometer wire plucked at the middle*

Factors affecting the frequency of a vibrating wire or string

The frequency produced by a vibrating string depends on the length of the string, l, and the velocity of the waves. On the other hand, the velocity of waves on a stretched string depends on the tension, T, in the string which is measured in newtons and the linear mass density, μ. The linear mass density is the mass of the string per unit length, that is,

$$\mu = \frac{m}{l} \text{ kg m}^{-1}$$

Therefore, the frequency of sound produced by a vibrating wire (string) depends on three factors; the tension (how taut the wire is), length and mass per unit length of the wire.

Activity 1.6

Aim: To investigate how the length of a vibrating wire affects its frequency.

Materials: Sonometer with a steel wire, a set of tuning forks, slotted masses, paper rider

Procedure:

1. Set up the sonometer as shown in Figure 1.40. The masses should be enough to keep the wire taut.

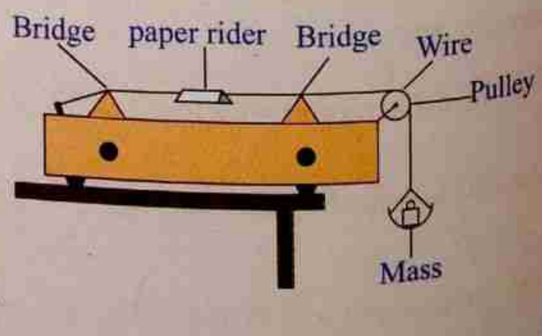


Figure 1.40

2. Set the paper ridder at the centre of the bridges.
3. Select a tuning fork with lowest frequency, set it into vibration and put it close to the paper rider but not in contact.
4. Adjust the bridges until the wire and the tuning fork reach the resonance (the paper rider flips off the wire).
5. Record the frequency, f of the tuning fork and the distance between the bridges.
6. Sound another tuning fork and without changing the mass, adjust the distance between bridges until the resonance is achieved.
7. Repeat the procedure for all the other tuning forks.
8. Record your results as shown in Table 1.3.

Table 1.3: *Results*

Frequency, f of the tuning fork	**Length, l of the wire**	$\frac{1}{l}$

Questions

(a) Draw a graph of f against $\frac{1}{l}$.

(b) Determine the gradient of the graph.

(c) From the graph, how does the frequency of a vibrating wire vary with increase in the length of the wire?

The graph of f against $\frac{1}{l}$ is a straight line through the origin. This means that frequency is inversely proportional to the length of the wire, that is, $f \propto \frac{1}{l}$.

On the other hand, if other factors remain constant, the frequency of the wire is directly proportional to the square root of tension, T, on the wire.

That is, $f \propto \sqrt{T}$

Furthermore, the frequency of the sound produced varies inversely with the mass per unit length, μ. If a thicker wire is used, the frequency decreases. From experiments, it has been established that frequency is inversely proportional to the square root of the mass per unit length.

That is $f \propto \sqrt{\frac{1}{\mu}}$

Note that, if the length of the vibrating wire is l, then, $l = \frac{\lambda n}{2}$

Recall that, $v = f\lambda$

so, $f = \frac{v}{\lambda} = \frac{nv}{2l}$

The velocity of a wave propagating through the string is given as $v = \sqrt{\frac{T}{\mu}}$

Thus, $f = \frac{n}{2l}\sqrt{\frac{T}{\mu}}$

For the first harmonic, $n=1$.

Thus, $f = \frac{1}{2l}\sqrt{\frac{T}{\mu}}$

This is the fundamental frequency of a vibrating string.

Since, $\mu = \frac{m}{l}$, then, $\mu = Area \times Density$

Therefore, $\mu = \pi r^2 \rho$ hence, $f = \frac{1}{2l}\sqrt{\frac{T}{\pi r^2 \rho}}$

Where r and ρ are respectively the radius and density of the string.

This equation shows how the fundamental frequency of a string depends on its length, l, tension, T in the string and its mass per unit length, μ. These results can be related to stringed instruments. For example, a guitar has six strings of the same length and these are held under approximately the same tension. However, the strings have different values of mass per unit length and so their fundamental frequencies are different. The larger the mass per unit length the lower the note and vice versa. Each of the strings is tuned by slightly varying the tension in the strings. The musician then plays different notes by pressing the strings against the frets on the fingerboard to vary the length of the strings.

Example 1.5

A string has a length of 75 cm and a mass of 8.2 g. The tension in the string is 18 N. What are the frequencies of the 1st and 3rd harmonics?

Solution

$$\mu = \frac{m}{l}$$

$$= \frac{0.0082\text{ kg}}{0.75\text{ m}}$$

$$= 0.011\text{ kg m}^{-1}$$

$$v = \sqrt{\frac{T}{\mu}}$$

$$= \sqrt{\frac{18\text{ N}}{0.011\text{ kg m}^{-1}}}$$

$$= \sqrt{\frac{18\text{kg m s}^{-2}}{0.011\text{ kg m}^{-1}}}$$

$$= 40.5\text{ m s}^{-1}$$

From,

$$f_n = \frac{nv}{2l}$$

For the first harmonic, $n = 1$,

then, $f_1 = \frac{v}{2l}$

$$f_1 = \frac{40.5\text{ m s}^{-1}}{2 \times 0.75\text{ m}}$$

$$= 27\text{ Hz}$$

For the third harmonic, $n = 3$,

then, $f_3 = \frac{3v}{2l}$

$$f_3 = \frac{3 \times 40.5\text{ m s}^{-1}}{2 \times 0.75\text{ m}}$$

$$f_3 = 81\text{ Hz}$$

Example 1.6

The vibrating length of a stretched wire is altered at constant tension until the wire oscillates in unison with a tuning fork of frequency 320 Hz. The length of the wire is again altered until it oscillates in unison with a tuning fork of unknown frequency. If the two lengths are 90 cm and 65.5 cm, respectively, determine the unknown frequency.

Solution

For constant tension,

$f \propto \frac{1}{l} \Rightarrow fl = \text{constant.}$

Therefore, $f_1 l_1 = f_2 l_2$

It follows that,

$$f_2 = \frac{f_1 l_1}{l_2}$$

Then,

$$f_2 = \frac{320\text{ Hz} \times 90\text{ cm}}{65.5\text{ cm}}$$

$$= 440\text{ Hz}$$

Forced vibrations and resonance

When a tuning fork is sounded and placed on a bench or a hollow box, the sound produced is quite loud and can be heard all over the room. The bench or box acts like an extended source (or many point sources) which are set into forced vibrations by the vibrating fork. These loud vibrations cause the sound to be louder but, since the rate of loss of energy is high, the sound dies off fast.

Forced vibrations are vibrations that occur in a system as a result of impulses received from a nearby vibrating system.

The response of the system that is set into forced vibrations is best when the driving frequency is equal to the natural frequency of the responding system. The responding system is then said to be in resonance with the driving frequency. Therefore, resonance occurs when a body or system is set into oscillation at its own natural frequency as a result of impulses received from some other system which is vibrating at the same frequency. A good example of resonance is when tuning a radio set to adjust the value of capacitance in a circuit until it has the same natural frequency of oscillation as that of the incoming signals.

Resonance in a closed pipe

If a tuning fork is sounded at the open end of a tube with the other end closed, the air in the tube resonates (vibrates freely) at a certain length of the tube. The resonance is observed as a loud sound produced in the tube when the proper length of air column is obtained. The first resonance occurs when air vibrates at its fundamental note or first harmonic. The fundamental note consists of one-quarter cycle as shown in Figure 1.41.

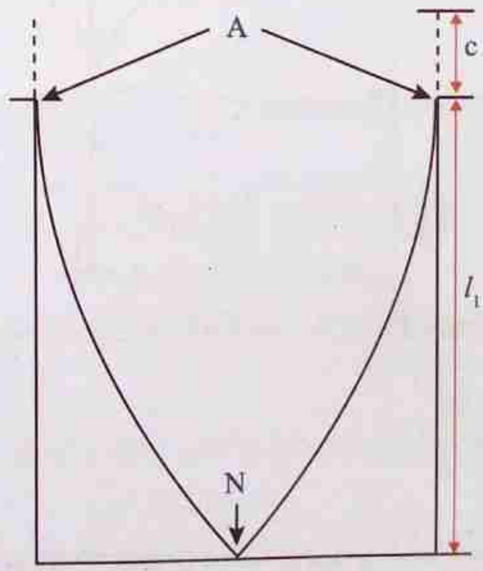


Figure 1.41: *The fundamental note (first harmonic)*

Note that, the length $l_1 = \frac{1}{4}\lambda$

In practice, the vibrations at an open end of a pipe extend into the free air, just outside. The actual position of the associated antinode is a short distance, c, beyond the end. This distance, c is known as the end correction. The effective length, l of a closed pipe is therefore, $l + c$ as shown Figure 1.41.

Considering the end correction, $l_1 + c = \frac{1}{4}\lambda$, where λ is the wavelength of sound wave.

The second harmonic or first overtone is produced when the length is increased to l_2 as illustrated in Figure 1.42.

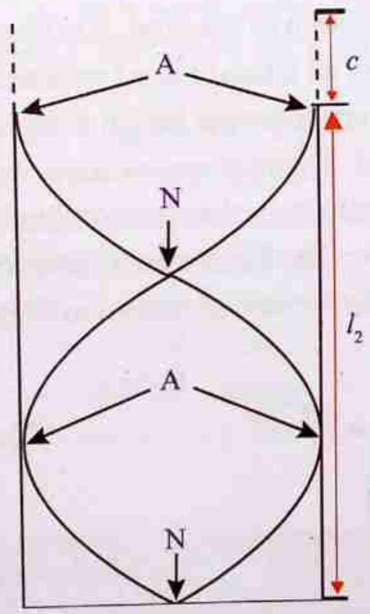


Figure 1.42: *Second harmonic (first overtone)*

$l_2 = \frac{3}{4}\lambda$ (neglecting end correction)

$l_2 + c = \frac{3}{4}\lambda$ (with end correction)

Increasing the length of the tube will give the third harmonic (Figure 1.43).

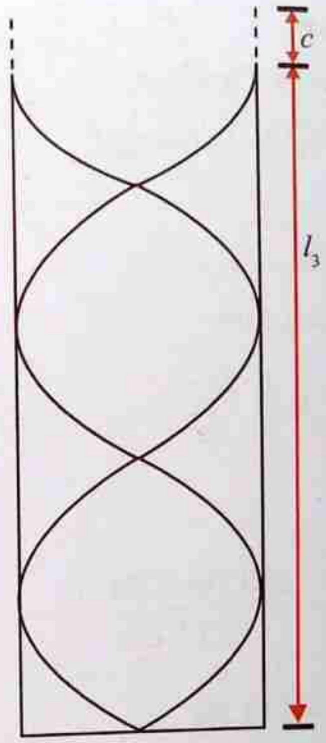


Figure 1.43: *Third harmonic (second overtone)*

$l_3 = \frac{5}{4}\lambda$ (without end correction), or

$l_3 + c = \frac{5}{4}\lambda$ (with end correction)

By subtracting l_1 from l_2 or l_3 from l_4, the end correction error can be eliminated. Generally, for any n^{th} overtone,

$$l_n = \frac{(2n+1)\lambda}{4} \text{ or}$$

$$l_n + c = \frac{(2n+1)\lambda}{4}$$

Note that, there is no way you can have even multiple of first harmonic in closed end pipes. This is because closed end pipes have a node at one end and antinode at the other end.

Activity 1.7

Aim: To investigate resonance in an air column.

Materials: Graduated cylinder, glass tube, thermometer, tuning fork, water

Procedure

1. Fill a graduated cylinder with water.
2. Partially insert a glass tube open at both ends.
3. Strike a tuning fork of known frequency and hold it at a position about 2 cm above the open end of the tube as shown in Figure 1.44.

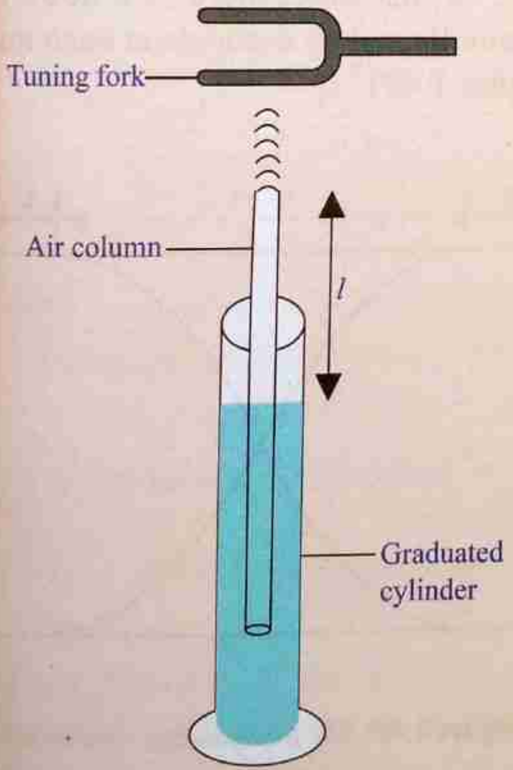


Figure 1.44

4. Raise or lower the tube in the water to vary the loudness of the sound produced in the air column.
5. Note the point at which the loudest sound is produced.
6. Determine the length 'l_1' of the air column of the tube at which the loudest sound is produced.
7. Repeat steps 3, 4, 5 and 6 to obtain a second loudest point and measure length l_2.
8. Measure the room temperature using a thermometer.

Questions

(a) Determine the wavelength of the sound produced.

(b) Calculate the speed of sound in air using the results obtained. Use the relationship: $v=2(l_2-l_1)f$

The water at the bottom of the tube acts as a seal, making the tube closed on one end. As the tube is raised or lowered, the volume of the produced sound changes. When the length of the air column produces a frequency that matches the frequency of the tuning fork, resonance will occur. As a result, the volume of the sound becomes relatively high.

Example 1.7

A tuning fork of frequency 512 Hz is sounded at the mouth of a tube closed at one end with a movable piston. It is found that resonance occurs when the column of air is 18 cm long and again when the column is 51 cm long. Use this information to determine the velocity of sound in air.

Solution

From $v = \lambda f$, and

$$c + l_1 = \frac{1}{4}\lambda$$

$$c + l_2 = \frac{3}{4}\lambda, \text{ we have:}$$

$$l_2 - l_1 = \frac{3}{4}\lambda - \frac{1}{4}\lambda = \frac{1}{2}\lambda$$

Where, $l_1 = 18$ cm and $l_1 = 51$ cm

$$\frac{1}{2}\lambda = l_2 - l_1 = (51-18) \text{ cm} = 33 \text{ cm} = 0.33 \text{ cm}$$

$$\lambda = 0.33 \text{ m} \times 2 = 0.66 \text{ m}$$

Therefore, $v = \lambda f = 0.66 \text{ m} \times 512 \text{ s}^{-1}$

$$v = 338 \text{ m s}^{-1}$$

Example 1.8

In a closed pipe, the first resonance is at 23 cm and second at 73 cm. Determine the wavelength of the sound and the end correction of the pipe.

Solution

From $l_1 + c = \frac{1}{4}\lambda$

Then, $0.23 \text{ m} + c = \frac{1}{4}\lambda$

Also, $l_2 + c = \frac{3}{4}\lambda$

Then, $0.73 \text{ m} + c = \frac{3}{4}\lambda$

$$(0.73 \text{ m} + c) - (0.23 \text{ m} + c) = \frac{3}{4}\lambda - \frac{1}{4}\lambda$$

$$0.50\ m = \frac{1}{2}\lambda$$

$$\lambda = 1 \text{ m}$$

From $0.23 \text{ m} + c = \frac{1}{4}\lambda$

$$0.23 \text{ m} + c = \frac{1}{4} \times 1 \text{ m}$$

$$0.23 \text{ m} + c = 0.25 \text{ m}$$

$$c = 0.25 \text{ m} - 0.23 \text{ m}$$

$$c = 0.02 \text{ m or } 2 \text{ cm}$$

Therefore, the wavelength of the sound is 1 m and the end correction is 2 cm

Resonance in open pipes

Consider the pipe which is open at both ends. In this case, there is a node in the middle and an antinode at each end (Figure 1.45).

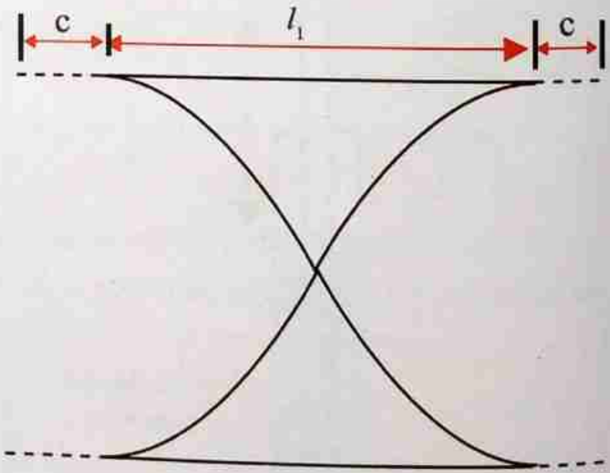


Figure 1.45: *The fundamental note in an open pipe*

The length of tube, l_1, is equal to half the wavelength, i.e.

$$l_1 = \frac{1}{2}\lambda \text{ (without end corrections).}$$

But the tube has two end corrections.

Therefore:

$$l_1 + 2c = \frac{1}{2}\lambda \text{ (with end corrections).}$$

By comparing the shortest lengths of the open pipe and the closed pipe, it is found that, the length of the open pipe is twice as much as that of the closed pipe. Note that, the pipes are resonating to a given frequency.

The second harmonic or first overtone is obtained by increasing the length of the tube (Figure 1.46).

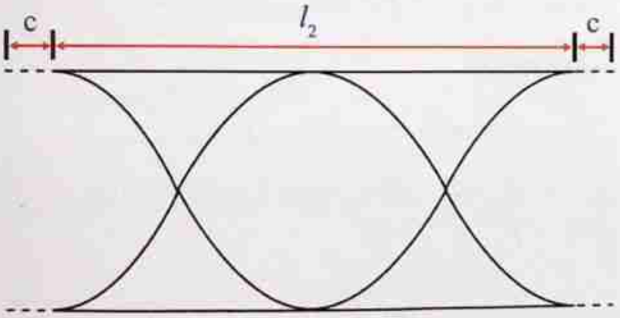


Figure 1.46: *Second harmonic in an open pipe*

$$l_2 = \lambda \text{ (without end corrections) or}$$

$$l_2 + 2c = \lambda \text{ (with end corrections)}$$

For any harmonic, n, the length l_n and wavelength are related by the following expressions:

$$l_n = \frac{n\lambda}{2} \text{ (without end corrections) and}$$

$$l_n + 2c = \frac{n\lambda}{2} \text{ (with end corrections).}$$

Example 1.9

A tuning fork of frequency 250 Hz is used to produce resonance in an open pipe. Given that the velocity of sound in air is 350 m s^{-1}, find the length of the tube which gives:

(a) the first resonance.

(b) third resonance.

Solution

From $l = \frac{\lambda}{2}$ for first resonance in open pipe;

then, $v = \lambda f$

$$\lambda = \frac{v}{f} = \frac{350 \text{ m s}^{-1}}{250 \text{ s}^{-1}} = 1.4 \text{ m}$$

(a) For the first resonance in open pipe,

$$l_1 = \frac{1}{2}\lambda$$

$$= \frac{1}{2} \times 1.4 \text{ m} = 0.7 \text{ m}$$

Therefore, the length of the tube in the first reasonance is 0.7 m

(b) For the third resonance,

$$l_3 = 3\frac{\lambda}{2}$$

$$= 3 \times 0.7 \text{ m} = 2.1 \text{ m}$$

Therefore, the length of the tube in the third reasonance is 2.1 m.

Beats

When to waves with slightly different frequencies travel in the same medium, a unique interference pattern appears. This happens when the two original waves combine constructively to form the largest amplitude or destructively to form the lowest amplitude as illustrated in Figure 1.47

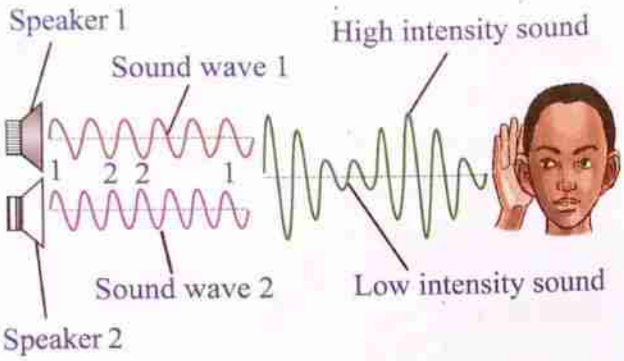


Figure 1.47: *Formation of beat frequency*

Suppose, two similar sound sources with nearly equal frequencies are sounded simultaneously, the oscillating loudness due to the oscillation of amplitude of the combined wave is heard. The sound rises and falls over time. This phenomenon is called beat.

Consider two tuning forks which are sounded with different frequencies. Let the first fork have 400 Hz and the second have 395 Hz.

Under these conditions they will produce a beat. The 400 Hz will have 40 cycles after $\frac{1}{10}$ s, while the 395 Hz tuning fork will have 39.5 cycles. Therefore, the 400 Hz tuning fork will be at compression, the 395 Hz will be at rarefaction. The resultant amplitude of the air will be minimum. Moreover, after $\frac{1}{5}$ s, a 400 Hz tuning fork will have 80 cycles while a 395 Hz will have 79 cycles. Therefore, at $\frac{1}{5}$ s the two waves are almost in phase. The resultant amplitude is maximum. In this case, the sound rises and falls after every $\frac{1}{5}$ s.

Therefore, there are 5 beats per second.

The beat frequency or the number of beats can be obtained by applying the principle of superposition given as the difference between the two frequencies of sound. Beat frequency $f = f_1 - f_2$ if $f_1 > f_2$ and $f = f_2 - f_1$ if $f_2 > f_1$

Example 1.10

A 256 Hz tuning fork produces sound at the same time with a 249 Hz fork. What is the beat frequency?

Solution

$$\text{Beat frequency} = f_2 - f_1$$
$$= 256 \text{ Hz} - 249 \text{ Hz}$$
$$= 7 \text{ Hz}$$

Exercise 1.3

1. A string of length 1.2 m is stretched and made to vibrate so that a stationary wave consisting of two loops is produced.
   (a) Draw a sketch of the wave.
   (b) Determine the wavelength of the wave.
2. A wire of length 20 cm, mass 1.2 g and under a tension of 120 N is plucked to generate a wave. Determine:
   (a) The fundamental frequency.
   (b) The frequency of the third harmonic.
3. The fundamental frequency of vibration of a string is f. What will

the fundamental frequency be if the length of the string is halved and the tension is increased four times?

4. A closed pipe has a fundamental frequency of 400 Hz. Calculate:
   (a) The frequency of the first overtone.
   (b) The fundamental frequency of an open pipe of the same length. (Given that the speed of sound waves in air, $v = 340\ \text{m s}^{-1}$)
5. A speaker delivering a note of frequency 250 Hz is placed over the upper end of a vertical tube filled with water. When the water is gradually run down the tube, the air column resonates when the water level is 31 cm below the top of the tube. The air column resonates again when the water level is 99.8 cm below the top of the tube. Determine:
   (a) The speed of sound in air.
   (b) The end correction.
6. If the shortest length of a tube for resonance is 0.12 m and the next resonant length is 0.37 m, what is the frequency of vibration? Take the speed of sound in air as $340\ \text{m s}^{-1}$.
7. A column of air 26.25 cm long in a closed tube resonates to a sounding tuning fork. If the velocity of sound in air is $33\,600\ \text{cm s}^{-1}$, what is the frequency of the fork? Ignore end correction.

Electromagnetic waves

Unlike mechanical waves which require a medium for their propagation, electromagnetic waves can propagate even in vacuum. These waves result from electric and magnetic fields that oscillate perpendicular to each other as illustrated in Figure 1.48.

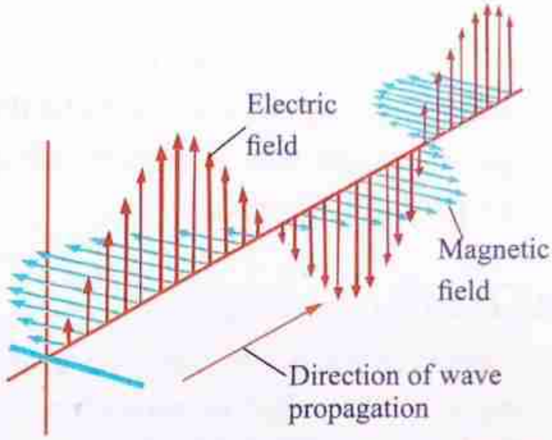


Figure 1.48: *Electromagnetic wave*

In order to initiate an electromagnetic wave, the motion of an electric charge must be altered. This is done by accelerating the charge. Thus, the sources of electromagnetic waves are accelerating charges. That is, neither stationary charges nor steady current can produce electromagnetic waves.

Electromagnetic waves are produced when electrically charged particles oscillate or their energies change. The larger the energy change, the higher the frequency of the resulting wave. In vacuum, electromagnetic waves propagate at the speed of light $3\times10^{8}\ \text{m s}^{-1}$. Examples of electromagnetic waves include radio waves, microwaves, infrared radiation, visible light, ultraviolet rays, X-rays and gamma rays. Visible light is the only electromagnetic wave that can be detected by the human eye.

Properties of electromagnetic waves

Electromagnetic waves are transverse waves which exhibit the following characteristics:

1. They can propagate in vacuum as well as in a material medium.
2. Electromagnetic waves undergo reflection, refraction, interference diffraction, scattering and polarisation.
3. All electromagnetic waves travel at the speed of light, that is approximately 3×10^8 m s^{-1} in vacuum.
4. They carry no electric charge.
5. They transfer energy from a source to a receiver in the form of oscillating electric and magnetic fields.
6. They obey the relation, $c = f\lambda$.
7. They can be polarised. Polarisation is the process that restrict electromagnetic waves to oscillate in a single plane or direction.

The electromagnetic spectrum

The electromagnetic spectrum is a continuous band of all electromagnetic waves arranged in the order of increasing wavelengths that is, decreasing frequency. A particular range of wavelength is called a band.

The electromagnetic spectrum is divided into seven major regions, namely: Radio waves, Microwaves, Infrared, Visible light, Ultraviolet light, X-rays and Gamma rays, as shown in Figure 1.49.

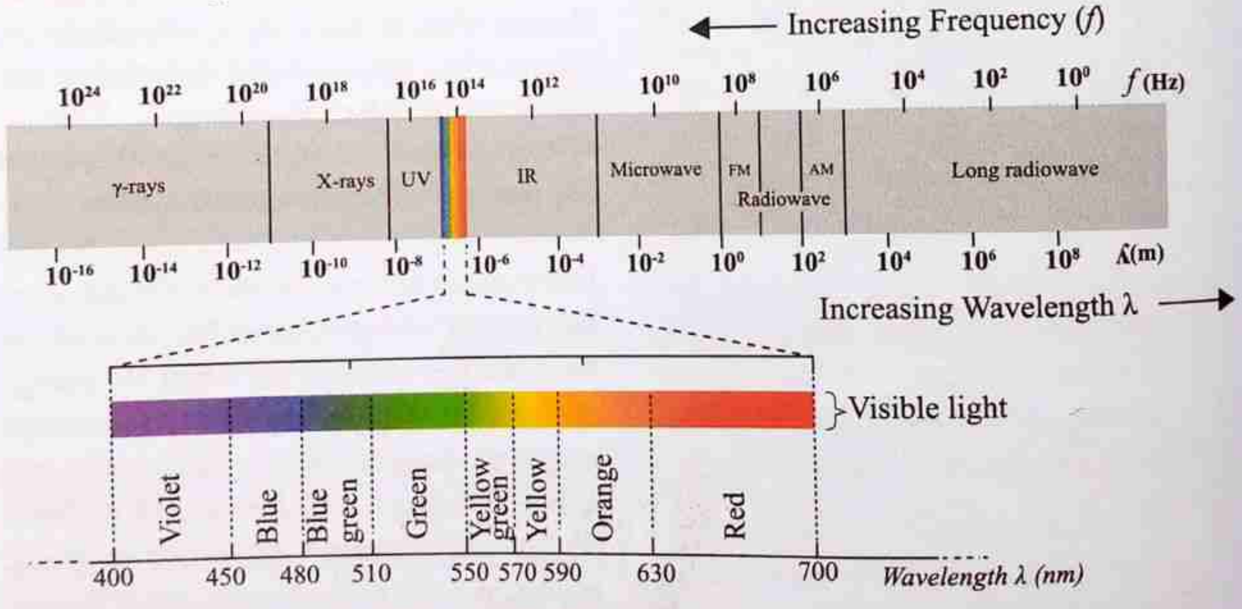


Figure 1.49: *Electromagnetic spectrum*

From Figure 1.49, the following observations can be made:

1. The electromagnetic spectrum is continuous. That is, each band merges into the next and there are no gaps in the frequencies. Different kinds of radiation gradually change from one kind to another as their properties gradually change.
2. In some cases, there is an overlap in the range of wavelengths. This is because sometimes the name given to the wave (radiation) is determined by the source of radiation and not the wavelength (or frequency), for example X-rays and gamma rays.

The range of frequencies and wavelengths for each band is as detailed in Table 1.4.

Table 1.4: *Range of frequencies of bands in electromagnetic spectrum*

Wavelength (m)	Region (band)	Frequency (Hz)
$>10^{-1}$	Radio waves	$> 3 \times 10^{9}$
$10^{-1} - 10^{-4}$	Microwaves	$3 \times 10^{9} - 3 \times 10^{12}$
$10^{-4} - 10^{-7}$	Infrared	$3 \times 10^{12} - 4.3 \times 10^{14}$
$7 \times 10^{-7} - 4 \times 10^{-7}$	Visible light	$4.3 \times 10^{14} - 7.5 \times 10^{14}$
$4 \times 10^{-7} - 10^{-9}$	Ultraviolet light	$7.5 \times 10^{14} - 3 \times 10^{17}$
$10^{-9} - 10^{-11}$	X-rays	$3 \times 10^{17} - 3 \times 10^{19}$
$< 10^{-11}$	Gamma rays	$> 3 \times 10^{19}$

Radio waves

Radio waves have the longest wavelength in the electromagnetic spectrum hence possess the least energy. Radio waves can further be divided into long waves (LW), medium waves (MW) and short waves (SW). Short waves include very high frequency (VHF) and ultrahigh frequency (UHF) waves.

Radio waves are generated by charged particles undergoing acceleration, such as time-varying electric currents. There are naturally occurring and artificially produced radio waves. Naturally occurring radio waves are emitted by astronomical objects in space such as planets, comets, stars and galaxies. Radio waves are generated artificially by transmitters and received by radio receivers, using antennas. Thus, radios are special devices used to transmit or receive radio waves.

Uses of radio waves

Radio waves have numerous advantages in human's daily activities. Along with other uses, the major applications of radio waves are found in communication. Some uses of radio waves include:

1. Radio waves are very widely used in modern technology for fixed and mobile radio communication, broadcasting, Radio Detection and Ranging (RADAR) and radio navigation systems, communications satellites, wireless computer networks and many other applications. Figure 1.50 shows a radio broadcasting station.

Figure 1.50: *Inside a radio broadcasting station*

2. Astronomers use large radio telescopes to collect and study radio waves from distant stars and galaxies. This helps them to determine composition, structure and motion of the celestial bodies. A radio telescope is shown in Figure 1.51.

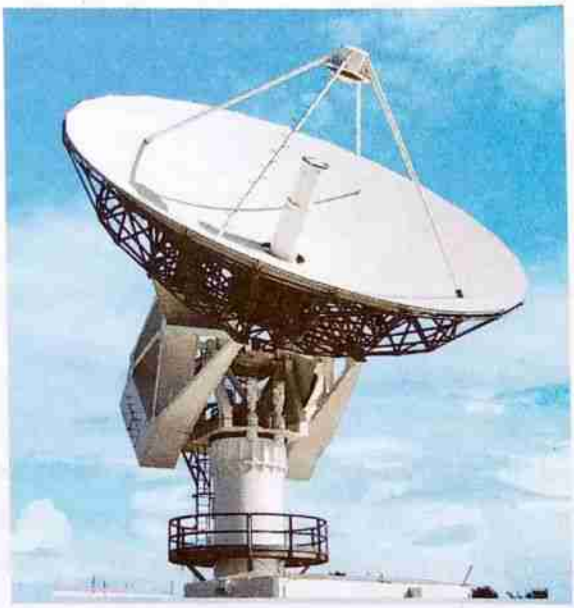


Figure 1.51: *Radio telescope*

Microwaves

Microwaves have wavelengths between 10^{-4} m and about 0.1 m. Their wavelengths are shorter compared to radio waves. Microwaves can be produced by artificial devices such as circuits, transmission towers, RADAR, and magnetron in microwave ovens. They can also be produced by natural sources such as stars and the Cosmic Microwave Background (CMB), a left over radiation after the Big Bang.

Uses of microwaves

Microwaves have many applications. Some of these applications are hereby described.

1. Microwaves are used in cooking. In this case, microwaves are absorbed by the food molecules. The absorbed energy causes the molecules to vibrate rapidly producing thermal energy that cooks or warms the food. Figure 1.52 shows a microwave oven.

Figure 1.52: *Microwave oven*

2. Some RADAR systems use microwaves to detect the position, speed, and other characteristics of remote objects such as aircraft and satellites.

3. Microwaves are used in long-distance communication because they are not affected by clouds or other atmospheric conditions. An example of a microwave transmitter is shown in Figure 1.53.

Figure 1.53: *Microwave transmitter*

Infrared waves

Infrared waves (infrared radiation) have wavelengths between 10^{-6} and 10^{-4} m. The band lies between the visible light and microwaves in the electromagnetic spectrum. Infrared waves near to the microwaves have a heating effect. Infrared waves are produced by the vibration of atoms and molecules due to their thermal energy. Almost all objects, including human bodies emit infrared waves.

Infrared radiation is invisible by the human eye but visible by nocturnal animals. However, humans can sense infrared radiation as heat. If you place your hand near an incandescent light bulb, you can feel the infrared radiation being emitted. Devices that are used to detect infrared radiation include black bulb thermometers, photographic films, thermistors and photo-transistors.

Uses of infrared waves

Infrared waves have several applications. These include:

1. Cooking or warming food in conventional ovens.
2. Infrared waves with wavelengths near the visible light are used in remote controls, night-vision devices, fibre-optic telecommunication and security systems.
3. Creating images in infrared photography. There are two techniques used to create images from infrared radiation, namely: Infrared photography and thermography.

Infrared photography uses a film that is sensitive to infrared radiation. For an object to produce an image on infrared, it must be at a temperature between 250 °C and 500 °C or reflect infrared radiation from a source in that temperature range. Infrared photography is used in long-distance photography because infrared is less affected by atmospheric haze compared to the visible light. Infrared photography is also used to detect the presence of disease in plants, and pollution in rivers and other water bodies. Besides, infrared photography is used in wildlife photography, particularly of nocturnal animals since it can produce images in almost total darkness. Figure 1.54 shows an infrared photograph.

Figure 1.54: *infrared photograph*

On the other hand, thermography or thermal imaging makes use of infrared receptors such as certain types of photoelectric and transistor cells. Thermography produces a thermogram which is colour-coded to highlight subtle temperature differences. Figure 1.55 shows a thermographic image of a hand. The following are some areas where thermography is used:

(a) Firefighters use it to see through smoke when finding fire victims, and locate the base of a fire.

(b) Electrical workers use thermography to locate overheating joints and parts in an electrical system.

(c) In medicine, thermal imaging is used to detect tumours and other diseased tissues, which tend to be warmer than the surrounding healthy tissue.

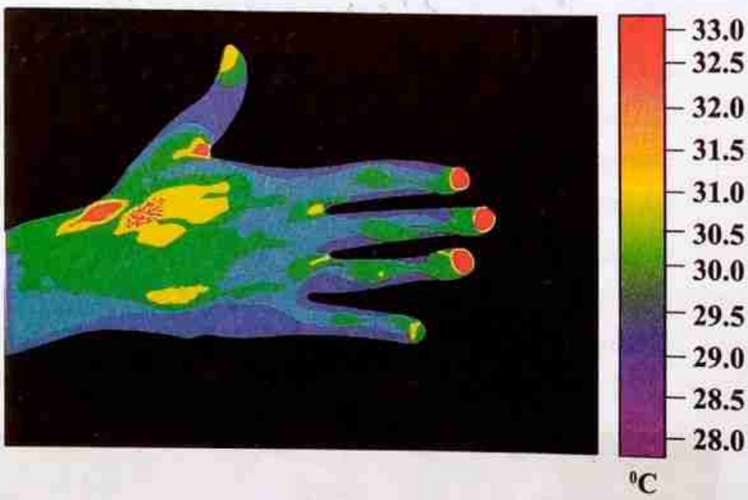


Figure 1.55: *Thermographic image of a hand*

Visible light

Visible light makes only a tiny fraction of the entire electromagnetic spectrum, yet it contains the frequencies to which human eyes are sensitive. The wavelengths that humans are typically able to visualize lie in a very narrow range between approximately 400 and 700 nanometres. Humans can observe and respond to stimuli created by visible light because the eyes contain specialized nerve endings that are sensitive to this range of frequencies.

Visible light can be produced by electron transitions within an atom. A vast majority of visible light is emitted from the Sun in a zone known as photosphere. Also, natural visible light is produced by celestial bodies, such as stars and galaxies. In addition, visible light originates from artificial sources, such fluorescent and incandescent tungsten devices. You see objects because they reflect visible light from any luminous source. For example, you see the Moon because it reflects light from the Sun. Note that, the Moon does not emit its own light. Visible light is detected using eyes, photographic films and photocells.

Uses of visible light

Visible light is very important to humans as it is used to see. Without light, our eyes would not be able to see images of anything. Beside seeing, there are other important uses of visible light, such as making lasers used as pointers. Visible light waves also make TV, computer and cell phone screens work. Visible light is used in photosynthesis by plants for making their own food. Visible light is also used by solar panels to generate electricity.

Ultraviolet radiation

Ultraviolet radiation has a shorter wavelength than visible light.

Like visible light, ultraviolet radiation is produced by electron transitions in atoms. However, ultraviolet light waves are more energetic than visible light. Ultraviolet radiation is also emitted by very hot objects. The Sun for example, emits ultraviolet radiation. However, most of the ultraviolet radiation from the sun is absorbed by the ozone layer in the atmosphere. Electric arcs used for welding also emit ultraviolet radiation.

Ultra violet radiation may be detected using:

1. Photographic films: Because ultraviolet light does not pass through ordinary glass, the camera must be fitted with quartz glass lenses.
2. Fluorescent materials, which absorb ultraviolet light and emit visible light as a result. Such materials glow upon being struck by ultraviolet light.

Uses of Ultraviolet radiation

The following are some of the uses of ultraviolet radiation:

(a) Ultraviolet radiation stimulates the production of vitamin D in the human skin. Vitamin D is essential in the prevention of certain types of cancer.

(b) Ultraviolet radiation is also used in the treatment of skin conditions such as psoriasis.

(c) Ultraviolet radiation is used as a germicidal agent in the sterilization of food and the purification of air and water.

(d) Fluorescent materials absorb ultraviolet radiation and emit visible light. Some washing powders contain fluorescent substances which glow in sunlight, making the clothes look brighter.

(e) Ultraviolet radiation is used in banks to detect forged documents and fake currencies.

(f) Ultraviolet radiation is used extensively in astronomical observatories to detect sources of young stars.

(g) Ultraviolet radiation is used in small industries such as electric welding.

Warning! *Prolonged exposure to ultraviolet radiation can lead to damages on the skin, eyes and the immune system.*

X-rays

X-rays are electromagnetic waves with short wavelengths and very high frequencies. X-rays are called ionising radiation because they can cause atoms and molecules with which they interact to lose electrons, thus producing ions.

X-rays can artificially be produced when highly accelerated electrons hit a metalic target. This process takes place in an X-ray tube.

X-rays can be detected using several devices including:

1. A photographic plate.
2. An X-ray film in a cassette.
3. Rare earth element screens.
4. Telescopes which detect X-rays emitted by astronomical objects such as stars and galaxies.

Uses of X-rays

X-rays have various applications in our daily life. These applications range from industrial to medical uses. Some of the applications of X-rays include:

(a) X-rays photography

When X-rays pass through a body, they are better absorbed by dense organs such as bone than by surrounding soft tissues. A photograph of the emerging X-rays displays a shadow of the denser organs on the film.

(b) X-rays are used in the diagnosis and treatment of cancer. Most of X-rays used in medicine are soft X-rays.

(c) X-ray diffraction is used to study the structural properties of materials.

(d) X-rays are also used for security checks at airport entrances and other sensitive areas.

Gamma rays

Gamma rays are the most energetic waves among the electromagnetic waves. They have the shortest wavelengths, thus, highest frequency. Like X-rays, gamma rays can cause ionization upon interaction with matter.

Gamma rays are produced in space by different sources such as solar flares, supernova, neutron stars, black holes and active galaxies. Nearly all gamma rays coming from space are absorbed by the Earth's atmosphere. On the Earth, gamma rays are produced by radioactive decay of atoms. The decay can be natural or artificial. Source of artificial radioactive decay include fission nuclear reactors, irradiators and atomic bombs.

Gamma rays can be detected using special instruments such as photographic films, Geiger–Müller tubes, cloud chamber, semiconductor detectors and space-based telescopes.

Uses of gamma rays

Gamma rays are the most penetrative radiation. They have many applications in the medical field especially as radiation therapy in diagnosis and treatment of cancer tumors. They are also used in agriculture to obtain new plant varieties which are disease-resistant and give more yields.

Similarly, gamma rays are used to destroy micro-organisms like fungi and bacteria. As a result, they are used for sterilization of medical-surgical equipment. Gamma rays are also used to sterilize food so as to improve hygiene using a technique known as food irradiation.

Exercise 1.4

1. Given the following electromagnetic radiation:
   (a) Infrared radiation
   (b) Ultraviolet radiation
   (c) Radio waves
   (d) Gamma rays

   (i) Which radiation has the longest wavelength?
   (ii) Which radiation has the highest frequency?

2. For each of the following electromagnetic waves describe a process by which it is produced.
   (a) Visible light
   (b) X-ray
   (c) Ultra violet
3. What is the speed of electromagnetic waves in a vacuum?
4. What is the origin of each of the following electromagnetic waves?
   (a) Radio waves
   (b) Infrared radiation
   (c) Visible light
   (d) Ultraviolet rays
   (e) X-rays

Chapter summary

1. Waves are disturbances that propagate from one point to another in a medium or in space. Waves transfer energy from one point to another.
2. Waves can be categorized as electromagnetic waves and mechanical waves. Waves can also be grouped into transverse waves and longitudinal waves depending on the mode of propagation.
3. Mechanical waves require a material medium to propagate and transfer energy while electromagnetic waves can travel through vacuum.
4. In transverse waves, the vibration of the medium particles is perpendicular to the direction of the wave propagation. In longitudinal waves, the vibration of the medium particles is parallel to the direction of the wave propagation.
5. The amplitude of a wave is the maximum displacement of the medium's particles from equilibrium.
6. The period of a wave is the time taken by the wave to make one complete cycle. The frequency of a wave is the number of cycles completed in a unit time. Wavelength is the distance a wave travels in one complete cycle.
7. Waves can undergo reflection and refraction at a boundary between two media.
8. Diffraction of waves is the apparent bending of waves around a small obstacle. It refers to the spreading out of waves when they pass through an aperture.
9. Wave interference is the addition or superposition of two or more waves propagating in the same medium. Superposition of waves results into constructive or destructive interference.
10. Sound is a longitudinal wave produced by vibrating objects. Sound waves are composed of variations in pressure. Sound travels fastest in solids and slowest in gases.
11. The ear converts sound energy to mechanical vibrations and then to electrical energy where the signal reaches the brain via nerves and is interpreted.

12. Sound is characterized by its pitch (frequency), loudness (amplitude) and quality (timbre).
13. A musical instrument is a device constructed or modified for the purpose of making music. Musical instruments are grouped into three broad categories, namely string instruments, percussion instruments and wind instruments.
14. Electromagnetic waves are waves which propagate through space or matter by the oscillation of an electric field and a magnetic field at right angles to one another.
15. The electromagnetic spectrum is a continuous band of all electromagnetic waves arranged in order of increasing wavelengths or decreasing frequencies.
16. The components of the electromagnetic spectrum are radio waves, microwaves, infrared radiation, visible light, ultraviolet radiation, X-rays and gamma rays. Each band has a unique set of characteristics and applications.

Revision exercise 1

Choose the most correct answer in items 1 to 3.

1. How does a transverse wave propagate in a medium?
   (a) Along the same direction as the direction of the particle vibrations.
   (b) From the displacement of the particle downward.
   (c) At right angle to the direction of the particle vibrations.
   (d) Towards the point where the particle was displaced.
2. Two waves come together and interact to form a new wave of smaller amplitude. What is the type of interference?
   (a) Destructive interference
   (b) Constructive interference
   (c) Reflective interference
   (d) Positive interference
3. Which of the following is the condition for wave refraction to take place?
   (a) Increase in velocity
   (b) Enter a new medium
   (c) Increase in frequency
   (d) Merge with another wave
4. Figure 1.56 shows the motion of a mechanical wave.

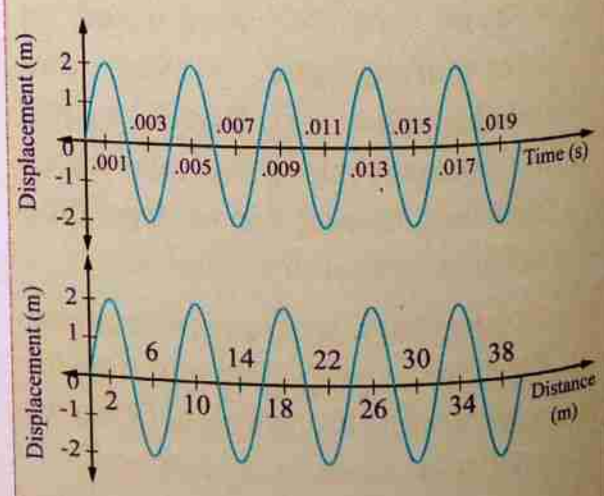


Figure 1.56

Determine the wave's:
   (a) Amplitude

(b) Period

(c) Frequency

(d) Wavelength

(e) Velocity.

5. Figure 1.57 shows water waves in a ripple tank moving from one region to another region. The two regions have different depths.

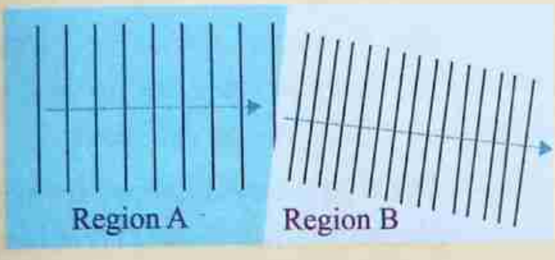


Figure 1.57

(a) In which region are the waves moving faster? Explain.

(b) Explain which region of the ripple tank is deeper.

6. During a storm, thunder is heard 7 s after the lightning is seen. If the temperature of the air at the time of the storm is 28 °C, how far away is the storm cloud?

7. A loud sound is made and an echo from a distant cliff is heard 8 s later. If the atmospheric temperature is 22 °C, how far away is the cliff?

8. A rope of length 80 cm and a mass of 10 g is set into vibration. If the tension in the rope is 20 N, find the frequencies of the 1st and 3rd harmonics.

9. Guitars have strings of varying thickness. Which of the strings, (thickest or thinnest), produce the highest frequency of musical notes? Explain your answer.

10. Matter expands when heated and contracts when cooled. Explain why a musician must re-tune a stringed instrument if its temperature changes.

11. Explain why it is not advisable for soldiers to march across a bridge in rhythm.

12. String A is 2 m long and has a linear mass density of 9 g cm^{-1}. String B has a linear mass density of 18 g cm^{-1}. If the tension in both strings is the same, how long must string B be for it to be at resonance with string A?

13. In a resonance tube experiment, the smallest value of l for which a peak in sound intensity occurs is 9.0 cm. How far must the tube be raised to hear the next peak in sound intensity? Neglect the end correction.

14. Figure 1.58 shows the main bands of the electromagnetic spectrum. Region D represents visible light.

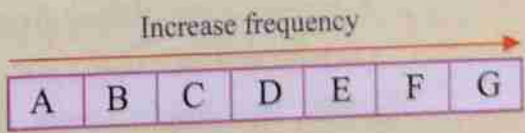


Figure 1.58

(a) Which region represents radiation capable of promoting the production of vitamin D in the skin?

(b) Which region contains radiation used in RADAR systems?

(c) Which region contains radiation produced in nuclear reactors?

15. About 50% of the radiation emitted by the Sun is visible light. Within this visible light, the wavelength that the Sun emits the most is approximately 500×10^{-9} m. What is the frequency of this light?

16. How long does it take for a radio signal sent from the Earth's surface to reach an object placed at a distance of 3.84×10^{6} m?

17. Consider two organ pipes A and B sounded with their:
    (i) Fundamental notes, and
    (ii) First overtones.

    If pipe A is closed at one end and pipe B is open at both ends, draw diagrams to show the mode of vibration in each of the four cases.

18. (a) How does the size of a gap in a barrier affect the diffraction of waves?
    (b) (i) State two ways in which visible light differs from radio waves.
    (ii) List two applications of gamma rays.

19. (a) What is meant by the following terms?
    (i) Resonance
    (ii) Overtones
    (b) Briefly, give reasons for the following:
    (i) The fundamental frequency may alter during the day.
    (ii) Notes of the same pitch played on a violin and a flute sound different.
    (c) The frequency obtained from a plucked string is 400 Hz when the tension is 2 N.

    Calculate:
    (i) The frequency when the tension is increased to 8 N.
    (ii) The tension needed to produce a note of frequency of 600 Hz.

20. A sonometer wire was tuned to a fundamental frequency of 250 Hz. The vibrating wire was then under tension of 10 N and was 87 cm long. Show by calculation how to tune the wire to a fundamental frequency of 375 Hz by adjusting:
    (a) Its length only.
    (b) Tension only.

Chapter Two

Electromagnetism

Introduction

In previous classes you learnt about magnetism as well as the electric current and its heating effects. Have you ever wondered about other effects of electric current and how they are related to magnetism? One of the effects of the electric current flowing through a conductor is the formation of magnetic fields. Electromagnetism is vital for designing and operating many devices. In this chapter, you will learn about concept of electromagnetism, magnetic field due to a current carrying conductor and electromagnetic induction. You will also learn about the working principles of electric devices such as induction coils, generators and transformers. The competencies developed will enable you to put into practice the principles of electromagnetism in designing and constructing simple electromagnets, electric motors, generators and transformers. You will also be able to perform scientific experiments related to electromagnetism and its effects.

Concept of electromagnetism

In 1820, a Danish professor, named Hans Christian Oersted, noticed that a compass needle lying near a wire was affected by the current flowing through the wire. Oersted was teaching his students a lesson on electricity with a demonstration involving a wire connected to a battery. When he closed a switch, an electric current started flowing through the wire and suddenly **a compass needle was deflected away from magnetic north.** From this observation, Oersted concluded that a conductor carrying an electric current is always surrounded by a magnetic field that is produced by the moving electrons. Oersted's discovery set the stage for our modern understanding of the interactions between electric fields and magnetic fields. Indeed, magnetic fields are caused by changing electric fields. Conversely, changing magnetic fields causes electric fields. Thus, electric current and magnetism are two related phenomena. That is to say electricity and magnetism are two aspects of electromagnetism.

Therefore,

A phenomenon associated with electric and magnetic fields and their interaction is known as electromagnetism. It is the effect produced by the interaction of an electric current with a magnetic field.

The interaction between electric and magnetic fields may result to a force causing the current carrying conductor to move. On the other hand, if a conductor is moved perpendicular to a magnetic field, a current is formed within the conductor.

Magnetic field due to a current carrying conductor

When an electric current flows through a conductor, it generates a magnetic field around the conductor. Thus, around any current carrying conductor, there is a magnetic field. This effect is demonstrated in Activity 2.1.

Activity 2.1

Aim: To demonstrate that an electric current flowing through a conductor produces a magnetic field.

Materials: Two dry cells, switch, rheostat, magnetic compass, ammeter, connecting wires, copper wire conductor

Procedure

1. Set up an electric circuit as shown in Figure 2.1.

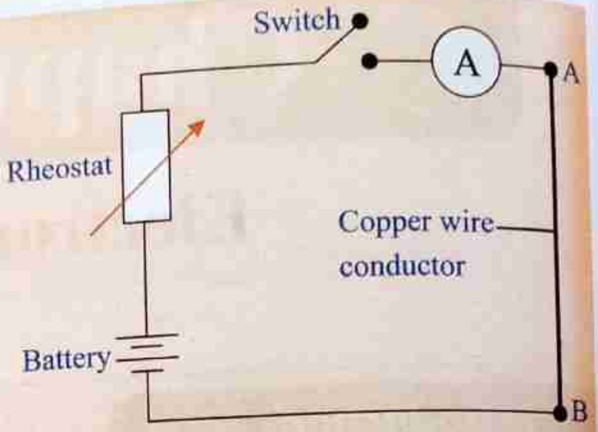


Figure 2.1

2. Hold a magnetic compass directly above conductor AB. Adjust the position of the conductor such that it is in line with the compass needle as shown in Figure 2.2.

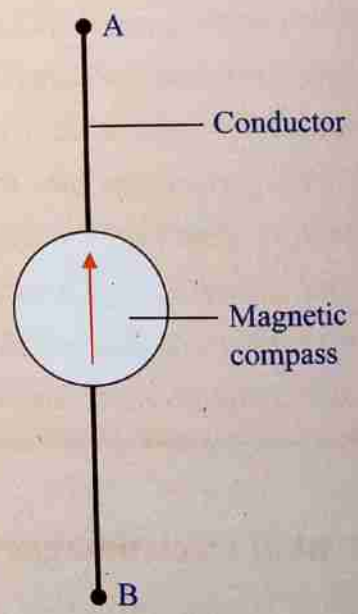


Figure 2.2

3. Close the switch to allow an electric current to flow through the circuit; and observe what happens to the compass needle. Record your observations.
4. Vary the current through the conductor AB. Observe the compass needle as the current through AB increases or decreases. Record your observations.

5. Take the compass and place it at about 5 cm and 10 cm above the conductor. In each case observe the needle. Record your observation.
6. Open the switch and place the compass below the conductor AB. Close the switch and observe the compass needle. Record your observations.
7. Reverse the battery connection and repeat step 2 to 6. In each case, record your new observations.

Questions

(a) What happens to the compass needle when the switch is closed?

(b) What happens to the compass needle when the current is increased or decreased?

(c) What happens when the compass gets farther from the wire?

When the switch is closed, an electric current flows through the conductor. The electric current generates a magnetic field around the conductor. This causes a deflection on the compass needle. The deflection gets smaller as the compass is drawn further from the current-carrying conductor. When the battery terminals are reversed, the compass needle deflects in the opposite direction. This means reversing the polarity results to a reverse in the direction of the produced magnetic field.

The magnetic field around a current-carrying conductor can be shown by the magnetic field lines or magnetic force as shown in Figure 2.3. These field lines can be visualized as a pattern of closed loops around the current carrying conductor.

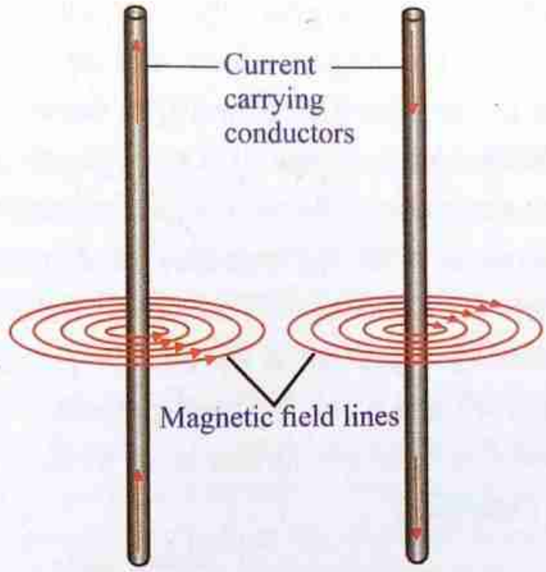
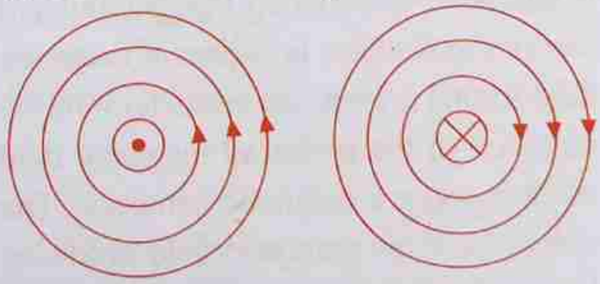


Figure 2.3: *Magnetic field lines around a current-carrying conductor*

The magnetic field pattern is usually described using a plane top view as shown in Figure 2.4. In the plane top view, the conductor is represented by a dot or cross. A dot in the circle in Figure 2.4 (a) indicates that, the current is flowing out of the plane (analogous to a tip of an arrow that is pointing out of the plane). On the other hand, a cross in Figure 2.4 (b), indicates that, the current is flowing into the plane (analogous to a tail of an arrow moving into the plane).



(a) Current out (b) Current in

Figure 2.4: *Plane view of the magnetic field pattern around a current carrying conductor*

The strength of the magnetic field produced by a current carrying conductor depends on the magnitude of the electric current. The higher the current, the stronger the magnetic field, and therefore the greater the deflection of the compass needle and vice versa. The strength of the magnetic field for a given current decreases with the increase in distance from the conductor. The strong field is normally represented by closely packed lines while a weak field is represented by field lines that are farther apart as shown in Figure 2.5.

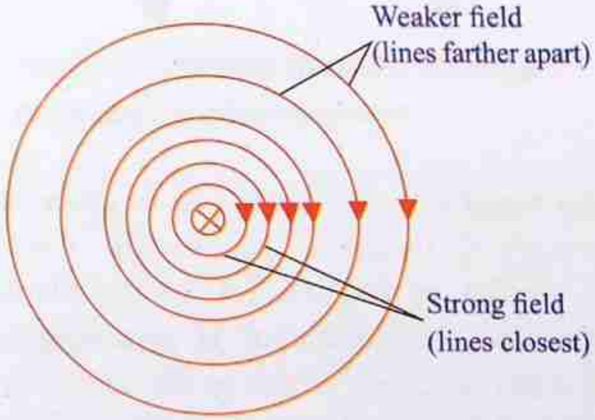


Figure 2.5: *Strength of the magnetic field around a conductor*

Direction of magnetic field produced by a current carrying conductor

It has been established that the current carrying conductor generates a magnetic field around it. How can one determine the direction of the produced magnetic field without using a magnetic compass? The direction of the magnetic field produced by a current carrying conductor can be determined by using some rules. These rules are the Right-hand grip rule and Maxwell's cork screw rule.

The Right-hand grip rule

The Right-hand grip rule can be applied to a straight conductor or a solenoid carrying an electric current. For a straight conductor, the Right-hand grip rule is as follows:

Imagine a conductor carrying current is gripped by the right hand with the thumb pointing in the direction of the conventional current. The fingers will curl around the conductor pointing in the direction of the magnetic field, around the conductor.

The Right-hand rule demonstrates the Oersted's discovery that, the direction of magnetic field changes when the direction of flow of electric current changes. Figure 2.6 shows the application of Right-hand grip rule for a straight conductor.

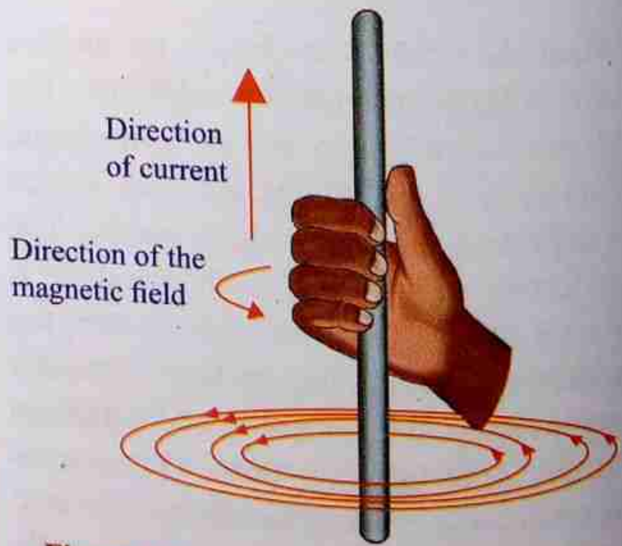


Figure 2.6: *Application of Right-hand grip rule for straight conductor*

The Right-hand grip rule can also be used for a current carrying solenoid. In this case, the rule is as follows:

When you wrap your right hand around a solenoid with your fingers pointing in the direction of conventional current, your thumb points in the direction of the magnetic North pole.

The magnetic field produced by a current carrying solenoid is similar to the magnetic field produced by a bar magnet and there are magnetic poles at the ends of the coil. The lines of magnetic force pass through the solenoid and return to the other end. Figure 2.7 shows the application of the right-hand grip rule for a solenoid.

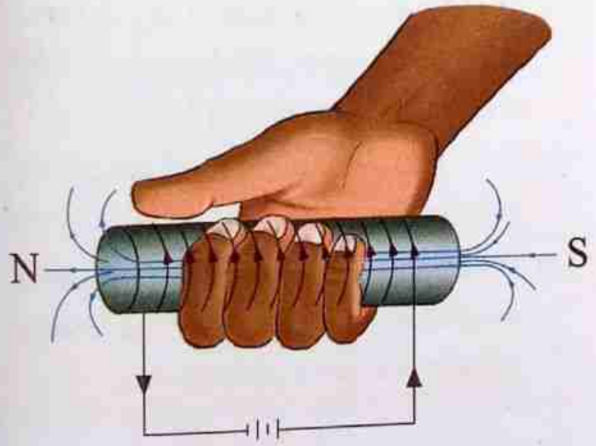


Figure 2.7: *Right-hand grip rule for a solenoid*

If a current-carrying solenoid is suspended freely, it comes to rest while pointing North (N) and South (S) like a suspended magnetic needle. One end of the solenoid acts like a N-pole and the other end a S-pole. Since the current in each circular turn of the solenoid flows in the same direction, the magnetic field produced by each turn of the solenoid adds up, giving a stronger resultant magnetic field inside the solenoid. The strength of the magnetic field in a solenoid can be increased by:

1. Making more turns of the coil,
2. Increasing the current, and
3. Inserting an iron core through the middle of the solenoid.

Maxwell's cork screw rule

Another rule that is used to determine the direction of the magnetic field produced by a current carrying conductor is the Maxwell's cork screw rule. Suppose you are driving a right handed screw into a piece of wood using a screw driver, there is a forward motion as the screw rotates in clockwise direction. Maxwell's cork screw rule considers the forward motion as the flow of current and the rotation as the direction of the magnetic field. The rule is as follows:

If a right-handed cork screw is rotated such that, its tip advances in the direction of the current flowing through a conductor, then the direction of rotation of the screw represents the direction of the magnetic field produced by the current.

Figure 2.8 demonstrates the Maxwell's cork screw rule.

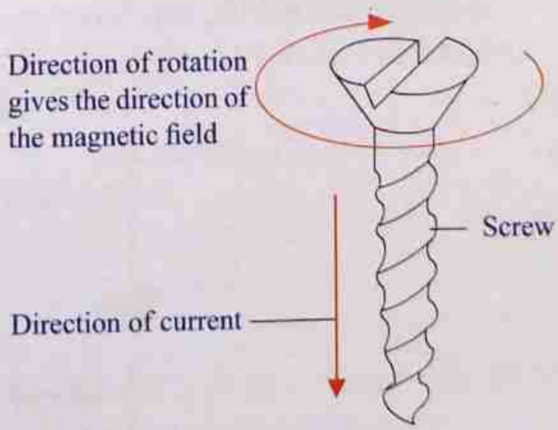


Figure 2.8: *Demonstration of Maxwell's cork screw rule*

Activity 2.2

Aim: To observe the magnetic field lines around a straight current-carrying conductor.

Materials: Accumulator, switch, rheostat, ammeter, connecting wires, conductor cardboard, iron filings and magnetic compass

Procedure

1. Connect an electric circuit with a conductor AB, placed vertically between points A and B as shown in Figure 2.9. Note that, conductor AB has to run through a sheet of cardboard.

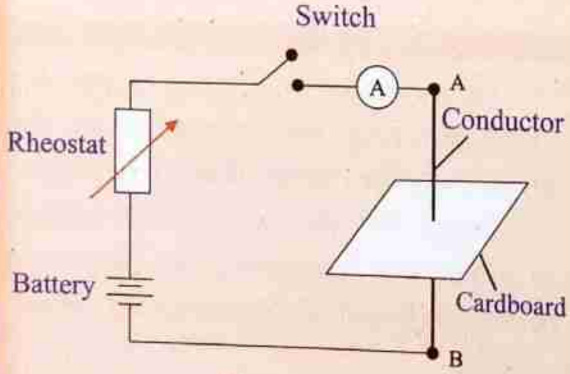


Figure 2.9

2. Sprinkle thin iron filings onto the cardboard around the conductor as shown in Figure 2.10.

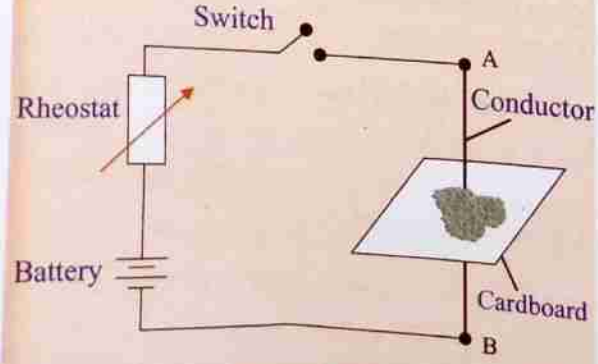


Figure 2.10

3. Close the switch so that electric current flows through the conductor.
4. Gently tap the cardboard a few times. Observe the pattern of the iron filings.
5. Hold a compass at various locations around the conductor and observe whether the magnetic field direction is clockwise or anticlockwise.
6. Reverse the battery connections and repeat step 2 to 5. Record your observations.

Questions

(a) What types of patterns are formed by the magnetic field produced by a current carrying conductor?

(b) How can the direction of magnetic field be found?

(c) What happens to the direction of the magnetic field lines when the battery connection is reversed?

When thin iron filings are sprinkled on the cardboard, and the current is switched on, the iron filings are set into concentric circles about the conductor at the centre. These circles represent magnetic field lines. If a compass is placed on the cardboard, it will indicate the direction of the magnetic field. If the battery connection is reversed, the compass will point in opposite direction but the pattern and the strength of the magnetic field remain the same. The direction of N-pole of the compass needle gives the direction of the magnetic field.

Task 2.1

Use Manila paper and coloured marker pens to neatly draw informative diagrams that illustrate the Right-hand grip rule and the Maxwell's cork screw rule. Present your diagrams in class.

Exercise 2.1

1. Give evidence of the existence of a magnetic field around a current carrying conductor.
2. Describe how you would use your right hand to determine the direction of a magnetic field around a current carrying conductor.
3. Write at least two applications of the Maxwell's cork screw rule.
4. In an experiment where an electric current flows through a conductor, a clockwise magnetic field is produced around the conductor. If a teacher decides to reverse the terminals of the battery, what would be the effect on the magnetic field?
5. Use the right-hand grip rule to find the direction of the magnetic field at each of the points labelled A to H in Figure 2.11.

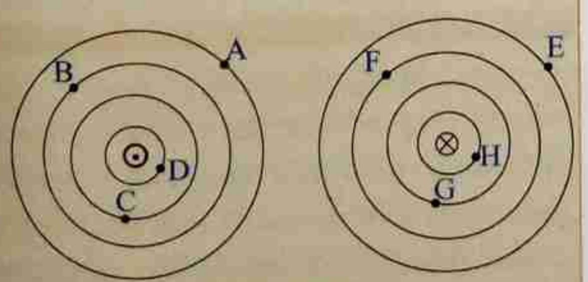


Figure 2.11

6. Briefly explain the factors which influence the magnetic field strength of a solenoid.

Force on a current-carrying conductor

It is known that, when an electric current flows through a conductor, it produces a magnetic field around the conductor. If a current carrying conductor is placed within an external magnetic field, the magnetic field around the conductor exerts a force on the external magnet field and vice versa. In turn there will be a force acting on the conductor. The direction of the force exerted on the conductor can be predicted using the Fleming's left-hand rule.

Activity 2.3

Aim: To investigate Fleming's left-hand rule.

Materials: Strong U-shaped magnet, two rigid copper rods (about 2 mm diameter and 20 cm long), one piece of rigid copper rod (5 cm long), wooden block of dimensions (10 cm × 10 cm × 5 cm), three dry cells, switch, connecting wires

Procedure

1. Bend the two 20 cm long copper rods at a point about 2 cm from one end to form L-shaped copper rods.
2. Fix the L-shaped copper rods on a wooden block and place a strong U-shaped magnet in such a way

that the two bent sides of the copper rods lie between the two poles of the U-shaped magnet as illustrated in Figure 2.12.

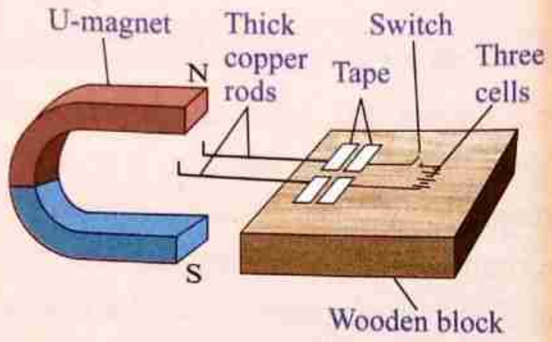


Figure 2.12

Note: The two rods should be separated by a distance of about 3 cm and run parallel to each other.

3. Bend the ends of the thin 5 cm long copper rod and complete the circuit by placing it at least 5 cm from the bent ends of the two L-shaped copper rods as illustrated in Figure 2.13.

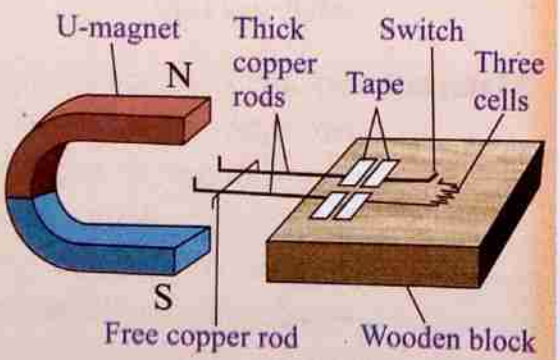


Figure 2.13

4. Close the switch and observe the motion of the free copper rod.
5. Open the switch and rotate the magnet through 180° about the horizontal axis so that the S-pole is on the upper side.
6. Close the switch and again observe the motion of the free copper rod.
7. Open the switch and return the magnet to its original orientation.
8. Reverse the direction of current by changing the orientation of cells.
9. Close the switch and observe what happens to the free copper rod. Record all your observations.

Questions

(a) Describe the movement of the free copper rod in each case.

(b) Why does the rod get displaced?

The displacement of the rod when the switch is closed suggests that a force is exerted on the conductor (5 cm long copper rod) which in this case is a current carrying conductor within the external magnetic field. The direction of this force is reversed when the terminals of the battery or the poles of the external magnet are reversed. Note that, the movement of the free rod is either towards the wooden block or away from it. This shows that the direction of the force on the conductor (5 cm copper rod) is perpendicular to the direction of both magnetic field and the electric current. That is, the force, the current and the magnetic field are mutually perpendicular to each other. This relation is summarized by the Fleming's left-hand rule which says;

If you hold the index finger, the middle finger and the thumb of your left hand such that they are mutually perpendicular to each other with the index finger pointing in the direction of the magnetic field and the middle *finger pointing in the direction of electric current, the thumb will point in the direction of the force acting on the conductor.*

The Fleming's left-hand rule is illustrated in Figure 2.14.

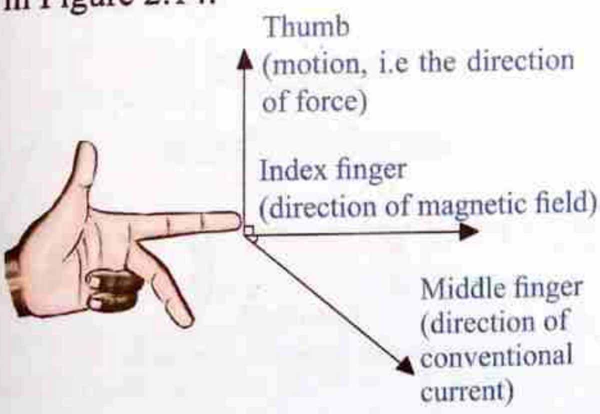


Figure 2.14: *Illustration of the Fleming's left-hand rule*

Force due to two parallel current carrying conductors

It has been established that, when a current carrying conductor is placed in an external magnetic field it experiences a force acting on it. Therefore, if two current carrying conductors are placed side by side, they tend to exert forces on each other due to their respective magnetic fields. The directions of these forces depend on the direction of currents through the conductors.

The forces on two parallel conductors can be investigated by performing Activity 2.4.

Activity 2.4

Aim: To determine the directions of forces produced by two parallel current-carrying conductors.

Materials: Two insulated and non-twistable wires (50 cm long), two accumulators, switch, connecting wires, crocodile clips, two smooth wooden blocks (10 cm × 5 cm × 5 cm) and thumb pins.

Procedure

1. Secure the smooth wooden blocks on the bench using adhesive tape. The distance from the end of one wooden block to the end of another block should be about 50 cm.
2. Stretch the wires between the wooden blocks and secure them in place using thumb pins.
3. Make a circuit as shown in Figure 2.15.

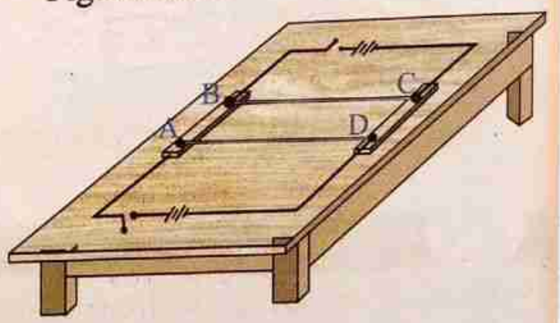


Figure 2.15

4. Close both switches and observe the behaviour of conductors AD and BC.
5. Open the switches and reverse the orientation of the terminals of the accumulator so that the current through conductors AD and BC are in the opposite direction. Observe the behaviours of the two conductors.

Questions

Explain the movements of the two wires when the currents are:

(a) In opposite directions.

(b) In the same direction.

When the currents flowing through the two parallel conductors are in the same direction, the two conductors attract each other, and when the currents flow in opposite direction the two conductors repel each other. Figure 2.16 illustrates the directions of forces between the two parallel current carrying conductors.

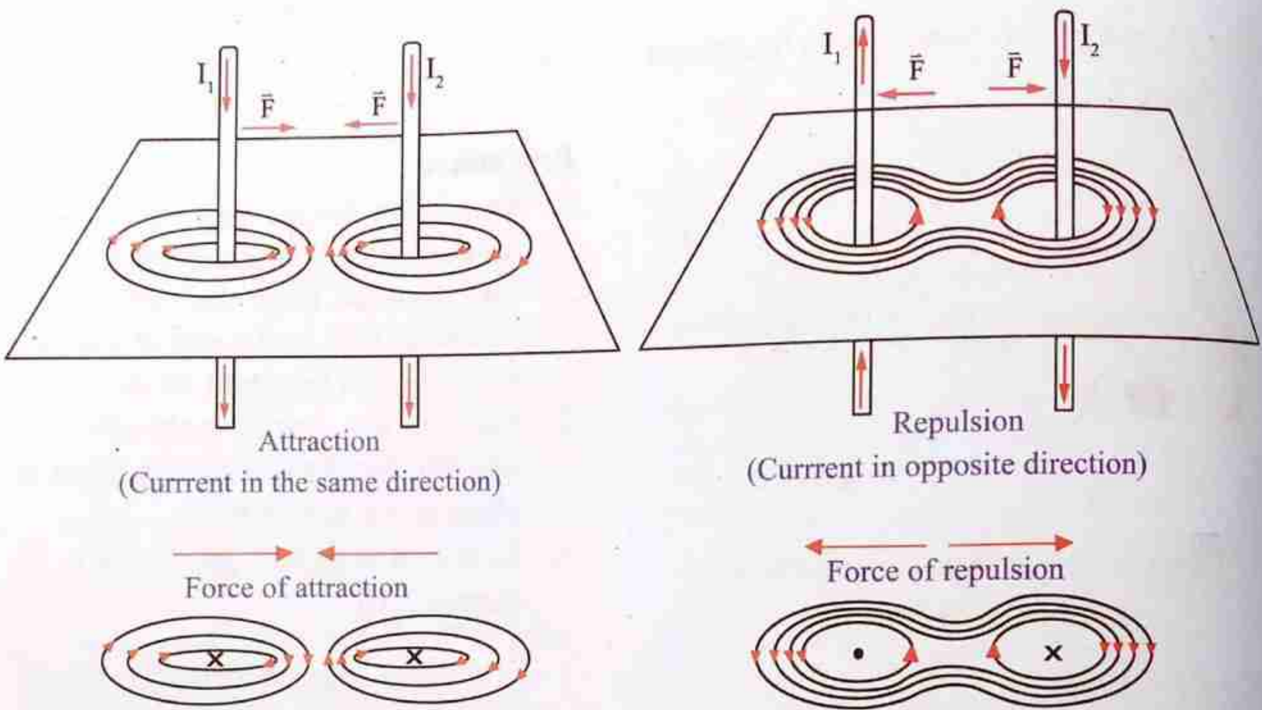


Figure 2.16: *Force between parallel current carrying conductors*

Task 2.2

You have been provided with connecting wires, source of power (dry cells), 2 aluminium strips of about 1 cm thick and 10 cm long. In groups, connect the circuit as shown in Figure 2.17.

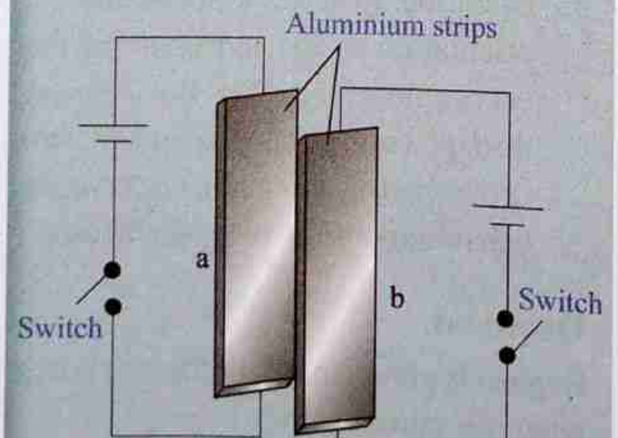


Figure 2. 17

Observe what happens to the aluminium strips when:

(a) Current is in the same direction and

(b) Current is in the opposite directions.

Present your observations in the class.

Electromagnetic induction

It has been seen that, a current flowing through a conductor creates a magnetic field around the conductor. One might ask if it is possible to create an electric current from the magnetic field. The answer to this question lies in the concept of electromagnetic induction. When a conductor is moved in a magnetic field such that the movement disturbs the magnetic flux through the conductor, an electromotive force (e.m.f) is generated across the conductor. The e.m.f generated leads to an induced current through the conductor. This phenomenon is known as electromagnetic induction. It is through this phenomenon that an electric current is formed from a magnetic field. The e.m.f and current produced through electromagnetic induction are respectively known as induced e.m.f and induced current.

Recall that, an electric current flowing through a conductor produces a magnetic field. On the other hand, moving a conductor inside a magnetic field or changing the magnetic field around a fixed conductor produces a current through the conductor.

When a conductor is moved across a magnetic field, a small e.m.f is generated in a conductor. If the conductor forms a complete circuit, the generated e.m.f makes current flow through the circuit. This current can be detected by deflection of a galvanometer connected in series with the conductor.

Activity 2.5

Aim: To investigate the induction of electric current through a conductor.

Materials: A strong horse-shoe magnet, a galvanometer, a bare conducting wire (about 30 cm long), wooden holder, two connecting wires and retort stand

Procedure

1. Connect the wire to the galvanometer to form the circuit as shown in Figure 2.18.
2. Clamp the horse-shoe magnet on a bench and set the wire-galvanometer circuit such that the bare conducting wire is between the poles of the magnet and approximately perpendicular to the magnetic field as shown in Figure 2.18.

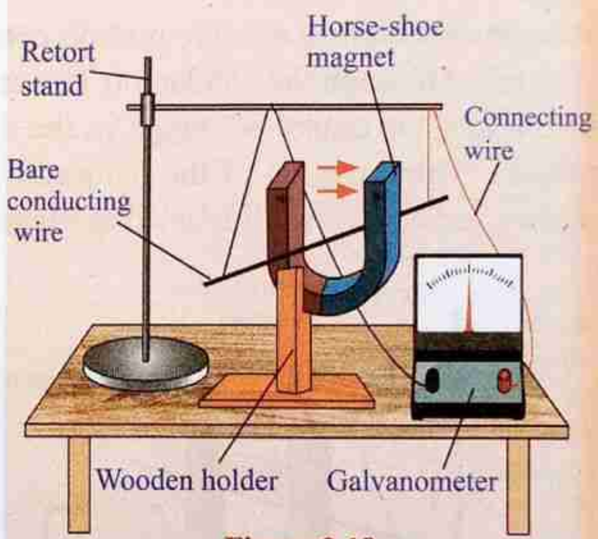


Figure 2.18

3. Hold the bare conducting wire and rapidly move it back and forth in a vertical direction between the poles. Observe the galvanometer as you move the wire.
4. Rapidly move the bare conducting wire back and forth in a horizontal direction between the poles. Observe what happens to the galvanometer.
5. Now, unclamp the magnet and fix the bare conducting wire between the poles of the magnet.
6. Hold the magnet and rapidly move it back and forth in a vertical direction. Observe the galvanometer.

7. Rapidly move the magnet back and forth in a horizontal direction. Observe what happens to the galvanometer.

Questions

(a) Explain what is observed on the galvanometer when:
   (i) The bare conducting wire was moved vertically between the poles of the magnet.
   (ii) The bare conducting wire was moved horizontally between the poles of the magnet.
   (iii) The magnet was moved vertically with the conductor between its poles.
   (iv) The magnet was moved horizontally with the conductor between its poles.

(b) Will the galvanometer deflect when both the bare conducting wire and the magnet are kept stationary?

Whenever there is a relative motion between the conductor and the magnet, current is induced through the conductor. This is because, the motion of either the conductor or the magnet causes a change in the magnetic flux linking the conductor and the magnet. The direction of the induced current is reversed when the direction of the motion is reversed as illustrated in Figure 2.19.

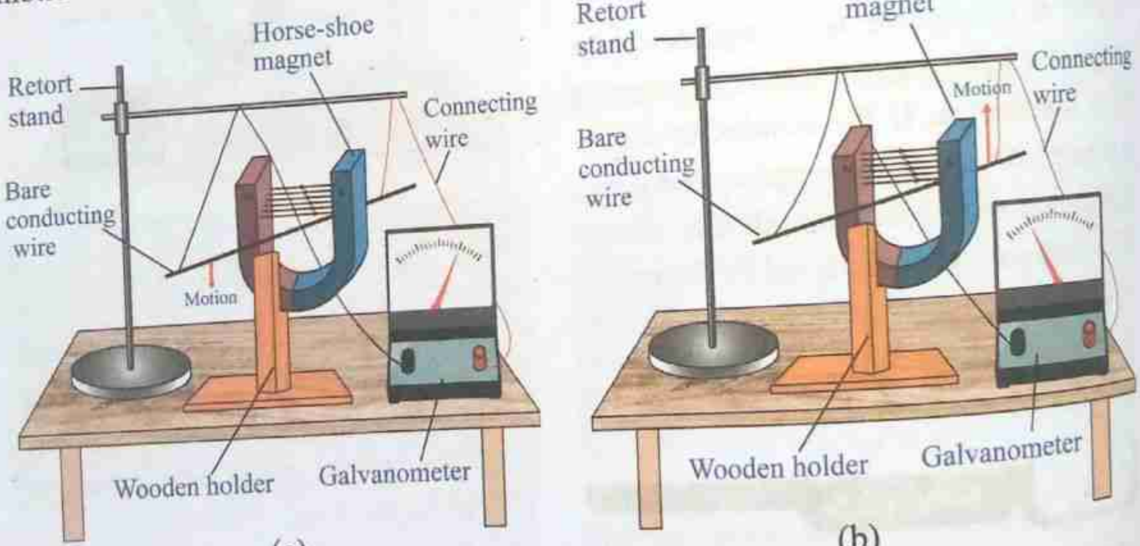


Figure 2.19

When the motion of the conductor is parallel to the field lines, no current is produced hence the galvanometer does not deflect as shown in Figure 2.20. This is because when the conductor is moved parallel to the direction of magnetic field, the magnetic flux linking the conductor and the magnet does not change.

That is to say, an induced e.m.f is produced only when the relative motion between the conductor and the magnet causes a change in magnetic flux linking the conductor and the magnet. The magnitude of the induced e.m.f increases as the speed of motion increase. There are two important laws that govern the phenomenon of electromagnetic induction.

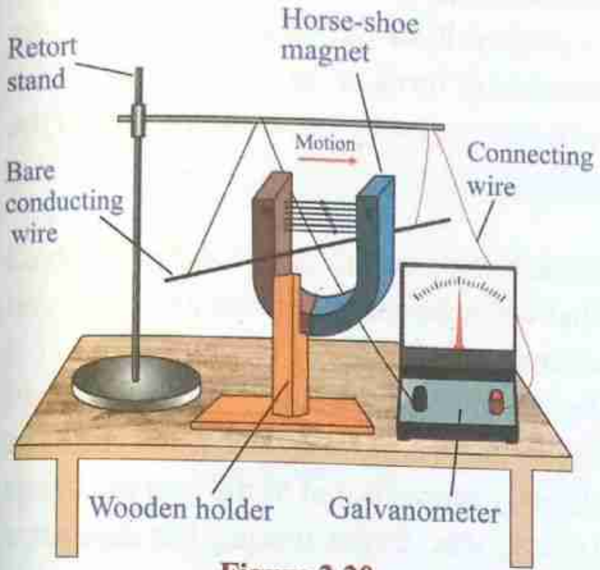


Figure 2.20

Laws of electromagnetic induction

To predict how the magnetic field would interact with a wire to produce an induced e.m.f, two laws have been established Faraday's law and Lenz's law. These laws allow physicists to describe the induction of e.m.f, direction or polarity, as well as the quantity of the induced current. These laws are collectively called the laws of electromagnetic induction.

Faraday's Law of electromagnetic induction

Faraday's law of electromagnetic induction is a basic law of electromagnetism that predicts how a magnetic field interacts with an electric circuit to produce an e.m.f. This law is fundamental in designing and construction of transformers, inductors, and many types of electrical motors, generators, and solenoids. The law relates the magnitude of the induced e.m.f and the rate of change of the magnetic flux linking the conductor and the external magnet. Faraday established this law after investigating various factors that affect the magnitude of the induced e.m.f.

Activity 2.6

Aim: To investigate the factors that affect the magnitude of induced e.m.f.

Materials: Two magnets (one strong, the other one weak), two coils of the same cross-sectional area but differ in number of turns (one with 2 000 turns and the other with 5 000 turns), two coils each of 4 000 turns but one with a larger cross-sectional area than the other, retort stand and clamp, centre-zero galvanometer, connecting wires

Procedure

1. Clamp the 2 000 turns coil horizontally about 30 cm above the bench.
2. Connect it in series with the galvanometer as shown in Figure 2.21.

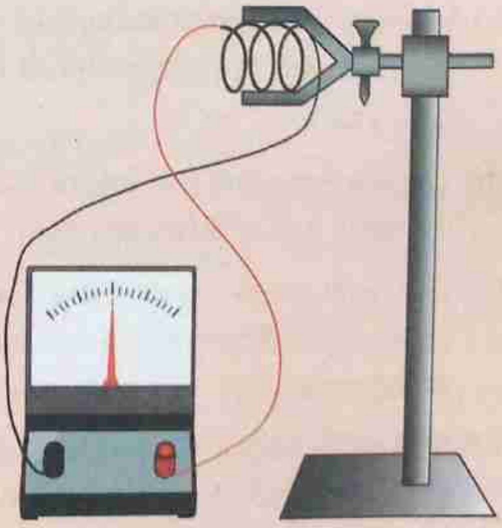


Figure 2.21

2. Insert the weaker magnet into the coil slowly while observing the galvanometer.
3. Repeat step 2, this time moving the magnet rapidly.
4. Repeat steps 2 and 3 using the stronger magnet.
5. Remove the 2 000 turns coil and replace it with the 5 000 turns coil of similar cross-sectional area.
6. Repeat steps 2 to 4 and record your new observation.
7. Replace the 5 000 turns coil with one of 4 000 turns of small cross-sectional area.
8. Rapidly insert the strong magnet into the coil. Record your observations.
9. Repeat step 8 using the 4 000 turns coil of a larger cross-sectional area.

Questions

How does the deflection of the galvanometer differ when:

(a) Magnets of different strengths were inserted into the same coil at the same rate?

(b) The same magnet was inserted into the same coil at different rates?

(c) The same magnet was inserted into coils with different number of turns at the same rate?

(d) The same magnet was inserted into coils with different cross-sectional area at the same rate?

The magnitude of the induced e.m.f depends on: the rate at which the magnet is pushed in or pulled out of the coil, the number of turns in the coil, strength of the external magnetic field and the area of the coil.

Based on these factors, Faraday observed that when the magnetic flux across the coil changes with respect to time, an e.m.f is induced across the coil. The magnitude of the induced e.m.f was also observed to be directly proportional to the rate of change of magnetic fluxes linking the conductor and the external magnet.

By generalizing these observations, Faraday come up with a law which states that:

Whenever the magnetic flux through a closed loop changes, an e,m,f is induced around the loop with a magnitude that is proportional to the rate of change of magnetic flux.

Lenz's law of electromagnetic induction

Faraday's law is useful in predicting only the magnitude of the induced e.m.f. Yet, it is important to predict also the direction of the induced e.m.f. Lenz's law was established to describe the direction of the induced e.m.f around the closed loop. If the induced e.m.f were to cause a current to flow through a conductor in an external circuit, that current would tend to generate a magnetic field that opposes the change in the original magnetic field. That is, the direction of the induced e.m.f is opposite to the change in the direction of the magnetic field. Activity 2.7 demonstrates the direction of the induced e.m.f.

Activity 2.7

Aim: Investigating the direction of the induced e.m.f.

Materials: Conducting wire, cardboard tube, galvanometer, bar magnet

Procedure

1. Make a coil using a conducting wire in the shape of a solenoid by wrapping several turns of the wire around a cardboard tube. Remove the cardboard tube and connect the coil to a galvanometer as shown in Figure 2.22.

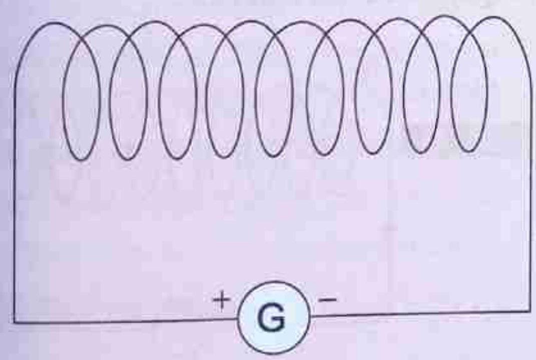


Figure 2.22

2. While observing the galvanometer pointer, bring the N-pole of the bar magnet towards the left end of the coil. Push the magnet about half its length inside the coil as shown in Figure 2.23. Observe the galvanometer while pushing the magnet.

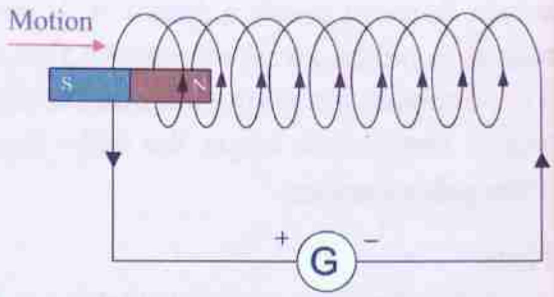


Figure 2.23

3. Slowly pull out the magnet until it is completely outside the coil while observing the galvanometer (Figure 2.24).

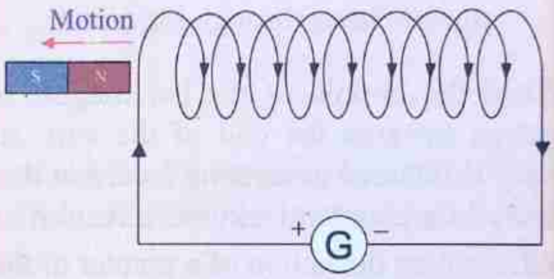


Figure 2.24

4. Repeat steps 2 and 3 using the S-pole of the magnet
5. Repeat steps 2 and 3 while holding the magnet stationary and moving the coil towards and away from the magnet.

Questions

In each case, describe the magnitude and direction of deflection of the galvanometer.

When the N-pole moves close to the coil, the magnetic flux through the coil increases. The galvanometer points toward the direction of the moving magnet. If a magnet is inserted into the coil with the N-pole on the front side, the current flows in positive direction. If the magnet is pulled away, the current flows in the opposite direction. The quicker the magnet moves, the larger the deflection of the galvanometer.

Note:

1. When the bar magnet is held fixed and the coil is moved towards or away from the magnet, the same effects are observed.
2. When there is no motion, there is no change in the magnetic flux and hence there is no induced e.m.f. That is, no current flows through the circuit.

When the N-pole of the bar magnet is moved towards the end of the coil, an e.m.f is induced generating a current that flows in a counterclockwise direction as indicated by deflection of a pointer of the galvanometer.

The induced current generates a magnetic field with its N-pole on the left side of the coil as shown in Figure 2.25 (a). However, when the N-pole is moved away or pulled out of the coil, the current flows in the clockwise direction and the poles of the coil are reversed (Figure 2.25 (b)). On the other hand, when the S-pole of a bar magnet is moved towards or away from the coil, the direction of the induced current in the coil is opposite to the one obtained when the N-pole of the bar magnet is moved towards or away from the coil. In this case, S-pole of the magnetic field of the coil is formed at the left side of the coil when the bar magnet is moving towards the coil. The left side of the coil becomes N-pole when the magnet moves away from the coil as illustrated in Figure 2.25 (c) and (d).

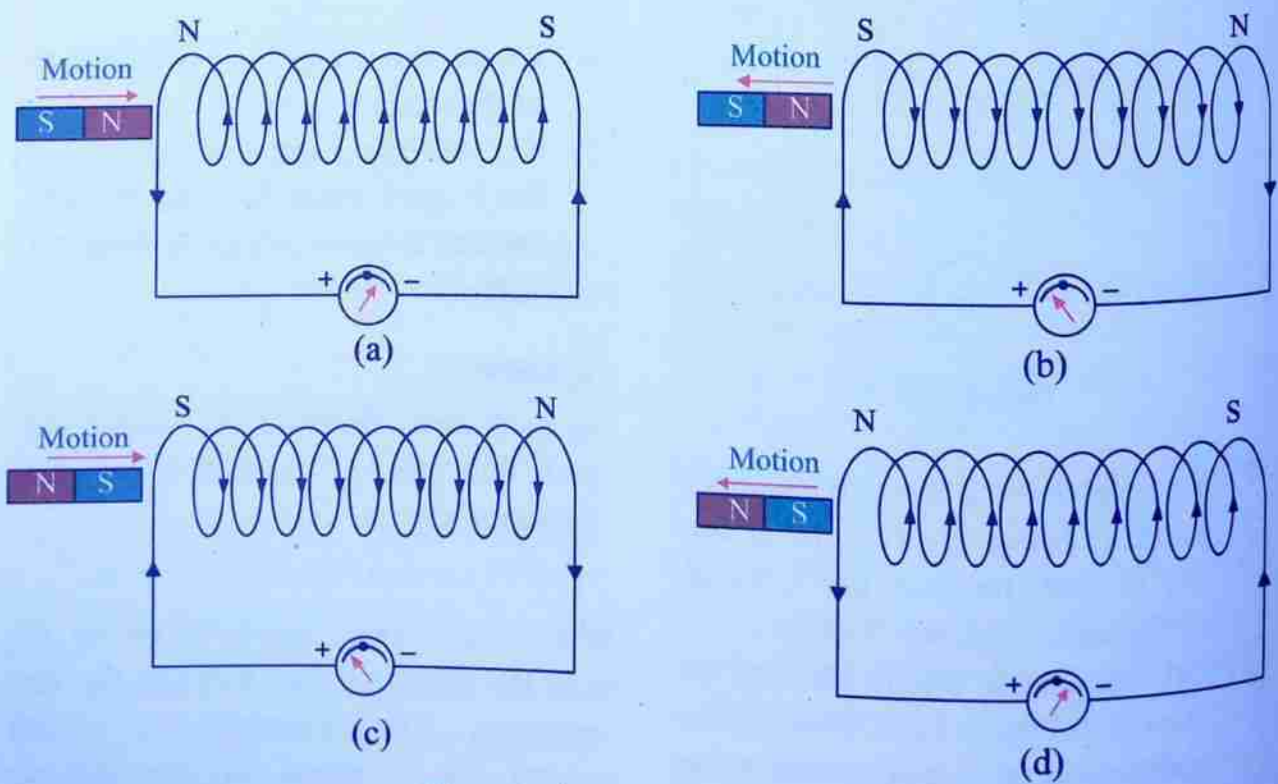


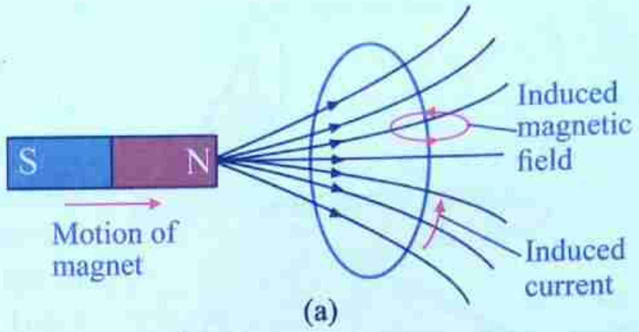
Figure 2.25: *Deducing Lenz's law*

These observations were first made by Emil Lenz who summarized them into a law known as Lenz's law of electromagnetic induction. The law states that:

The polarity of the induced e.m.f is in such a way that it tends to produce a current which opposes the change in the magnetic flux that produces it.

In other words, the induced current will produce a secondary magnetic field that opposes the change in the magnetic flux of the primary field. When the N-pole of a magnet is moving into a coil, the magnetic flux increases, this induces an e.m.f in the coil. The resulting induced current must be in a direction so as to oppose this increase in flux. The induced (secondary) magnetic field must also be oriented such that it opposes the motion that causes the increase in magnetic flux as demonstrated in Figure 2.26 (a).

Conversely, when the N-pole of a magnet is moving out of the coil, the magnetic flux decreases, which induces an e.m.f in the coil. The resulting induced current must be in the direction so as to oppose this decrease in flux. The induced (secondary) magnetic field must again be oriented to oppose the motion of the magnet as demonstrated by Figure 2.26 (b).



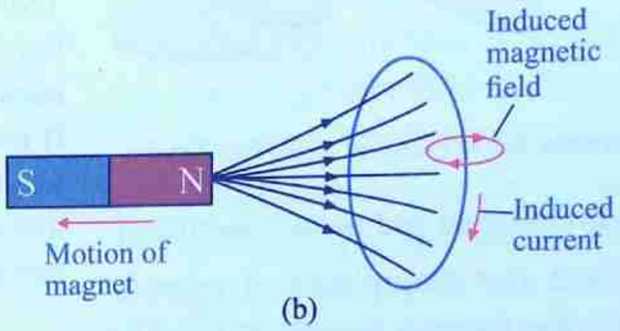


Figure 2.26 (a) and (b): *Direction of the induced current with respect to the direction of motion*

Determining the direction of the induced current requires a rule that was established by J.A. Fleming.

Fleming's right hand rule

Fleming could determine the direction of the induced current using his right hand. He thus established a rule which is referred to as *Fleming's right hand rule* or *dynamo rule*. The rule states that:

If the thumb, first finger and second finger of the right hand are extended mutually at right angles such that the first finger points towards the direction of the field and the thumb towards the direction of the motion of a conductor relative to the magnet, then the second finger will point towards the direction of induced current.

Application of Fleming right hand rule is pictorially described by Figure 2.27.

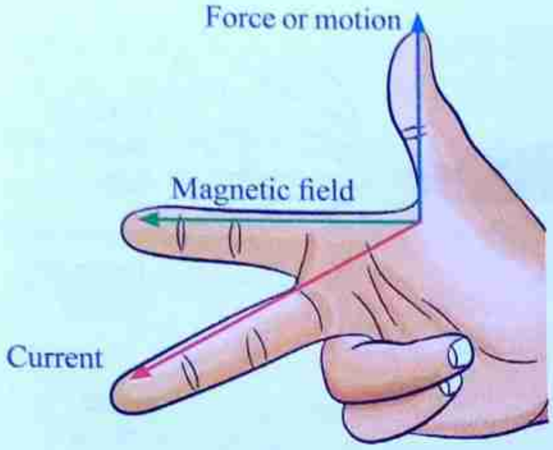


Figure 2.27: *Fleming right hand rules*

Self-induction and mutual induction

When the magnitude of current flowing through a conductor (coil) varies it creates a varying magnetic field that cuts across the conductor itself. The varying magnetic field induces an e.m.f in the same coil. The induced e.m.f will tend to oppose the variation of the magnetic flux across the coil. This phenomenon is known as *self-induction*. It refers to the induction of an e.m.f in a current carrying conductor as a result of variations of the magnitude of current flowing through the conductor itself. The induced e.m.f acts just like a second voltage source and is referred to as *back e.m.f.* That is, if the original current is increasing, the induced current subtracts from the original current and thus the measured current becomes smaller than it would be in the absence of the back e.m.f. Conversely, if the original current is decreasing, the induced current adds to it and the measured current becomes larger than it would be in the absence of the back e.m.f.

Consider a coil of wire wrapped around a cardboard tube making two loops. When the coil is connected to a battery, electric current flows producing a magnetic field (see Figure 2.28). If the current in the coil begins to increase, the magnetic field changes resulting to the change of magnetic flux. This induces an e.m.f that opposes the change in magnetic flux. That is, the induced e.m.f opposes the increase in the e.m.f produced by the battery resulting to a back current that blocks the increase in current. If the current in the coil begins to decrease, the magnetic flux decreases. This induces an e.m.f that adds to the battery resulting in an induced current that prevents the decrease in current.

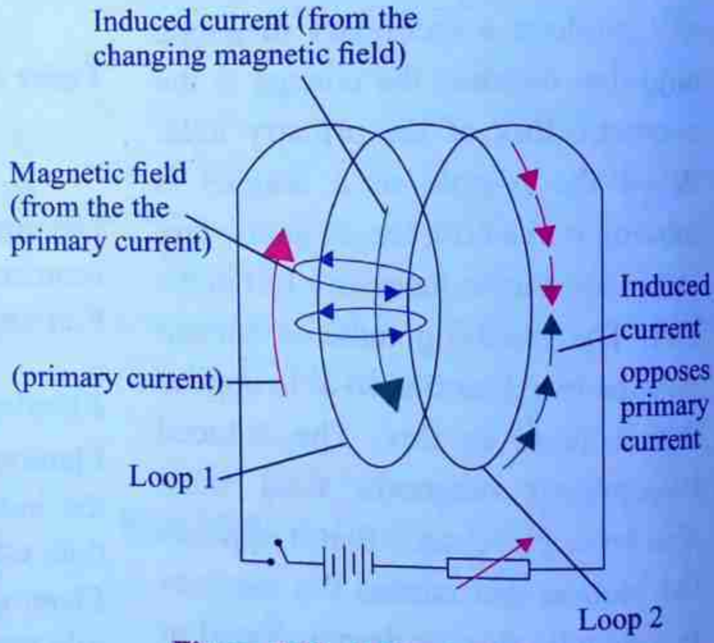


Figure 2.28: *Self induction*

If two coils are placed near each other, as shown in Figure 2.29, a varying current in one coil will induce a current in the other. This is called *mutual induction*. The coil with a changing current is referred to as the primary coil while that in which a current is induced is the secondary coil. If the current in the primary coil is not varying, there will be no e.m.f induced in the secondary coil. This is because, the magnetic field B_1 produced in the primary coil pass through the secondary coil since the coils are placed close to each other.

Therefore, when the current in the primary coil is changing, the magnetic field, through the secondary coil also changes. This produces a change in the magnetic flux in the secondary coil. In response, an e.m.f is induced in the secondary coil producing a current that in turn produces a magnetic field, B_2, that tends to oppose the change in the magnetic flux, B_1.

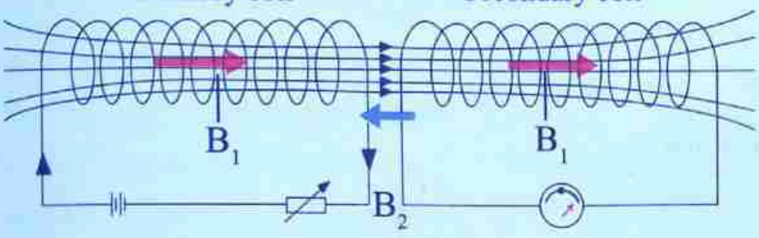


Figure 2.29: *Mutual induction*

The e.m.f induced in the secondary coil is proportional to the rate of change of current in the primary coil. The phenomenon of mutual induction has many applications in real life. These include the induction coil, generator and transformer.

Induction coil

An induction coil (Figure 2.30) is an electrical device consisting of two coils of insulated wires wound around a common iron core. The *primary coil,* is made from relatively few (tens or hundreds) turns of coarse wire. The *secondary coil*, typically consists of a larger number (thousands or millions) of turns of a fine wire. The secondary coil is normally wound on top of the primary coil. An induction coil is used to produce high voltage alternating current from low voltage direct current.

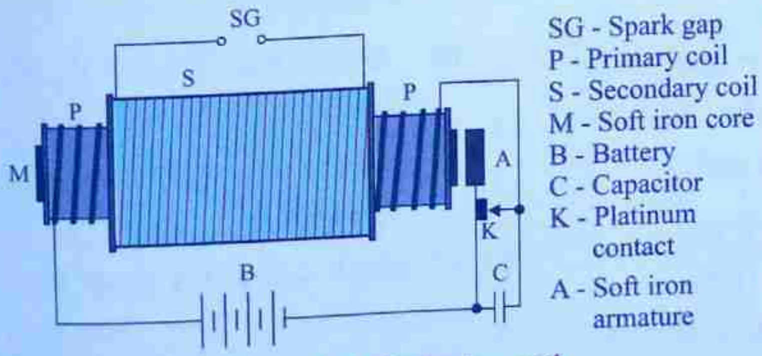


Figure 2.30: *Induction coil*

Mode of action

An induction coil produces high voltage in its secondary coil by electromagnetic induction. The direct current in the primary coil is switched on and off by a make-and-break mechanism. This produces changes in current and magnetic field which are necessary for electromagnetic induction to occur in the secondary coil. When the current in the primary coil is switched on, the induced magnetism in the iron core attracts the soft-iron armature. The moving armature causes the platinum contact to separate and opens a gap between the two contacts which breaks the primary coil circuit. This switches off the current. As the induced magnetism fades, the armature springs back, thus closing the contacts that completes the circuit again. This allows the current to flow in the primary coil again. This cycle of events is repeated automatically and an e.m.f is induced in the secondary coil. The induced e.m.f is very large, usually in the order of hundreds of kilovolts (kV) causing a spark to jump across the gap. Such a high voltage is achieved because of two aspects:

1. The secondary coil is having a large number of turns compared to the primary coil.
2. The current in the primary coil is switched on and off rapidly causing a rapid change in the magnetic field through the secondary coil.

Applications of the induction coil

(a) An induction coil is commonly used in ignition systems of internal combustion engines.

(b) A smaller version of an induction coil is used to trigger flash bulb tubes used in cameras and strobe lights.

(c) An induction coil is also used in wireless telegraphy.

(d) An induction coil is used in cooking, example induction cooker.

Exercise 2.2

1. Discuss the factors that determine the magnitude of the induced e.m.f in a closed loop of wire.
2. In Faraday's experiments, what would be the advantage of using coils with many turns?
3. Briefly explain why a galvanometer connected to a coil deflect when a magnet is moved towards the coil.
4. Two circular coils X and Y are placed close to each other. If the current in coil X is changed, will some current be induced in the coil Y? Explain.
5. A stiff wire AB is held between the poles of two permanent bar magnets and connected to a galvanometer with flexible wire as shown in Figure 2.31.

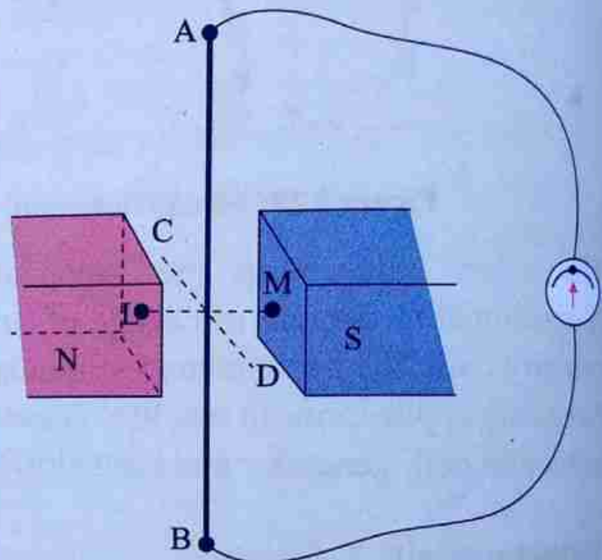


Figure 2. 31

The wire AB is kept vertical and moved slowly back and forth along the line CD.

(a) Describe and explain the behaviour of the galvanometer pointer:

(i) As the wire AB is moved towards C;

(ii) When the wire AB is momentarily at rest; and

(iii) As the wire AB is moved towards D.

(b) In which direction does the induced current flow when the wire AB is moving towards C?

Generators

An e.m.f is induced in a conductor placed within a magnetic field whenever there is a change in the magnetic flux linking the conductor and the magnet. According to Lenz's Law, the generated e.m.f opposes the change in magnetic flux. Therefore, in order to produce usable induced e.m.f, work must be done to overcome the resistance to change in magnetic flux. This is done by some special devices such as a generator. This device produces electricity on the basis of electromagnetic induction by the continuous motion of either a coil or a magnet. There are two types of generators; an alternating current (a.c) generator and a direct current (d.c) generator.

The a.c generator

An a.c generator is designed based on Faraday's Law of electromagnetic induction. It works by rotating a coil at a constant rate in a magnetic field to induce an oscillating e.m.f. An a.c generator consists of a source of magnetic field, an armature, slip rings, brushes and a resistive load as shown in Figure 2.32. The magnetic field is usually produced by a permanent magnet. An armature is made up of several turns of insulated wire wound on a soft-iron core. The armature revolves freely on an axis between the poles of a powerful magnet. Two slip rings are connected to the ends of the armature and two carbon brushes rest on the slip rings to conduct current from the armature to a load.

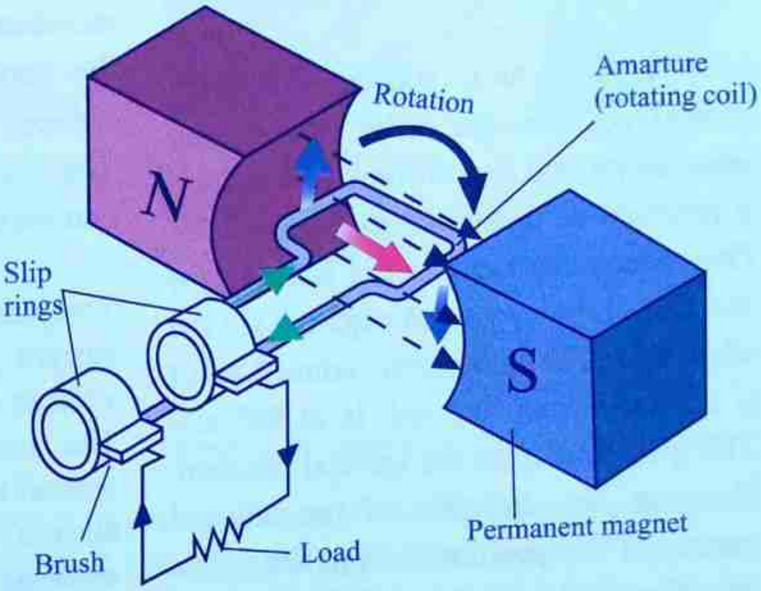


Figure 2.32 *Parts of a simple a.c generator*

When the coil is vertical as shown in Figure 2.33, no cutting of the magnetic lines of force takes place. The rate of change of magnetic flux is zero, thus, no e.m.f is induced in the coil.

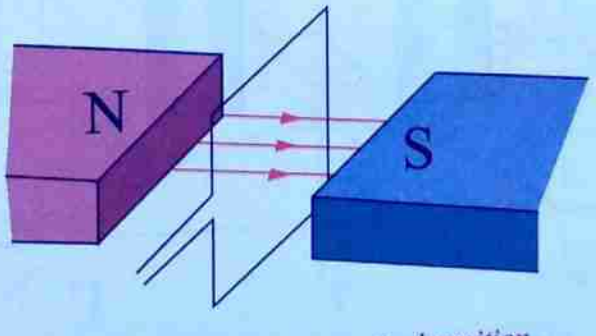


Figure 2.33: *Coil at vertical position*

When the coil is horizontal as shown in Figure 2.34, the rate of change of magnetic flux is maximum and hence the motion of the coil results to an induced e.m.f in the coil.

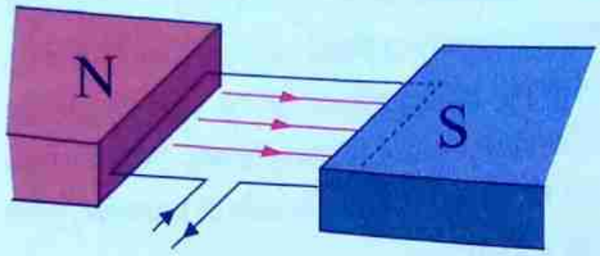


Figure 2.34: *Coil at the horizontal position*

After a 180° turn, starting from the vertical position, the sides of the loop interchange and the current in the loop is reversed as illustrated in Figure 2.35. This means that the e.m.f is positive for one half of the cycle and negative for the other half. The maximum induced e.m.f is obtained when the coil is at 90° and 270° positions from the vertical position. However, the direction of the induced current at 90° position is opposite to the direction of the current at 270° position.

If there is an external circuit, the current through it would also have a maximum value at 90° and 270°.

Therefore, the induced current starts from zero at the vertical position and continuously increase to the maximum value at 90° position. It then starts to decrease to zero at 180° and then increases in opposite direction towards the maximum at 270° position. It then reduces to zero at 360° (vertical) position. The cycle then repeats as the coil continues to rotate.

This kind of current is called an alternating current (a.c) and the corresponding voltage (e.m.f) is an alternating voltage. The number of cycles produced per second is called the frequency of the a.c current. The a.c obtained is tapped by an external circuit through the slip rings and the carbon brushes.

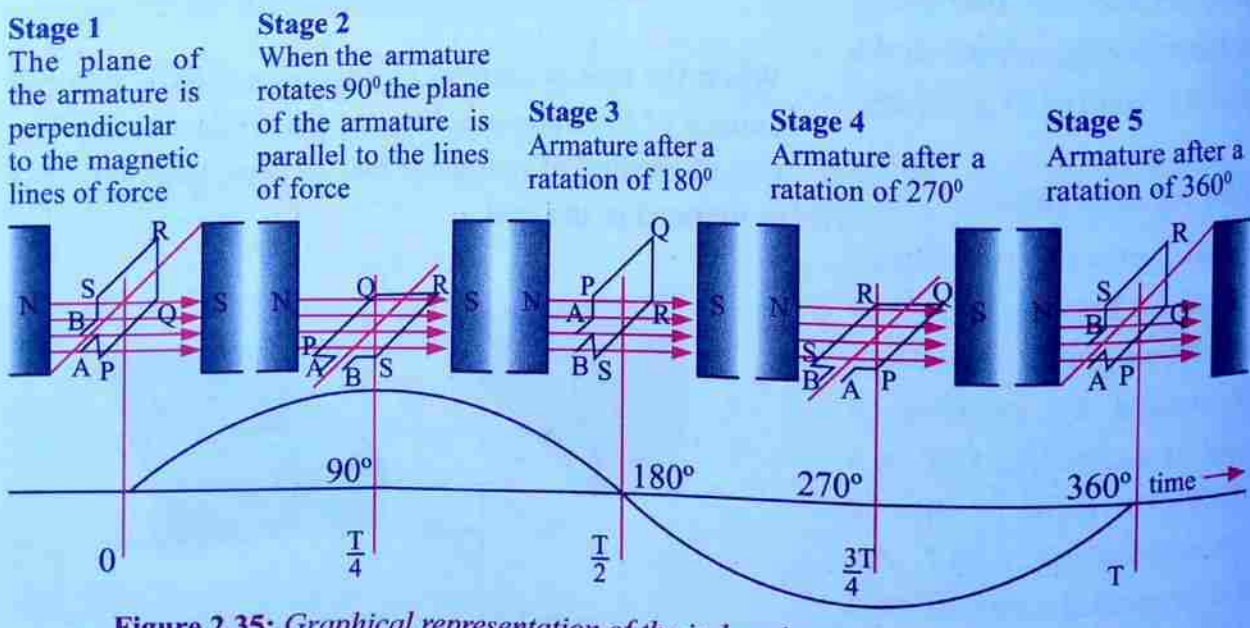


Figure 2.35: *Graphical representation of the induced current in the a.c generator*

Most electric generators are designed to produce a.c. These generators are often called alternators. Practical alternators usually have a fixed set of coils arranged around a rotating electromagnet. Most cars, for example, are fitted with alternators which produce electricity for running the car after it is ignited. The produced current (a.c) is changed into direct current (d.c) using diodes. Part of this d.c is used for charging the car battery. For a car alternator, the armature is rotated by a belt which is moved by the crankshaft of the engine. In power stations, much larger alternators are used to generate a.c electricity. The energy needed to rotate the such large armatures in power stations is obtained from sources like falling water from a dam, compressed steam from a nuclear reactor, or steam from underground. The basic principle of generating electricity in all alternators is the same.

The d.c generator

A d.c generator is made by replacing the slip rings of the a.c generator with a commutator. Each half of a commutator ring is called a commutator segment and is insulated from the other half. Each end of the rotating loop of wire is connected to a commutator segment. Two carbon brushes connected to the outside circuit rest against the rotating commutator. The commutator rotates with the loop of wire, just as the slip rings do in the a.c generator. Figure 2.36 shows a simple d.c generator.

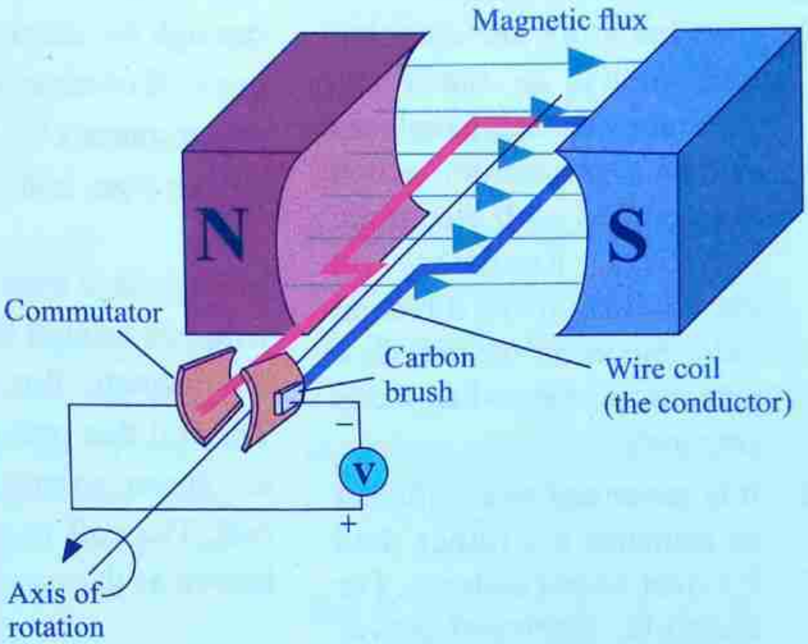


Figure 2.36: *Simple d.c. generator*

When the loop is rotated in the magnetic field, the induced e.m.f is still an alternating e.m.f. However, after a rotation of 180°, instead of the current reversing, the connections to the external circuit are reversed so that the direction of current direction in the external circuit remains the same. The output of a d.c generator is as shown in Figure 2.37.

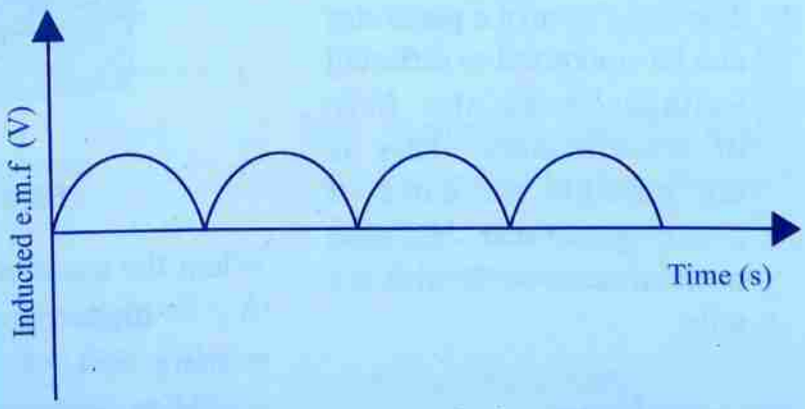


Figure 2.37: *Output of a d.c generator*

Note that, the lower half of the cycle is not cut off but reversed.

Advantages of a.c over d.c generators

There are several advantages of using a.c generators over d.c generators. These include:

1. Commutators are complex and costly, so many d.c generators are being replaced with a.c generators which are coupled with electronic rectifiers. Rectifiers let current flow in one direction only, thus allow the use of a.c generators even when d.c is required.
2. It is easier and more efficient to transmit a.c rather than d.c over a long distance. For example, electrical power is often generated many kilometres from where it is ultimately used. This kind of electrical power transmission can easily be achieved with a.c electricity.
3. The design of the a.c generator is simpler than the design of a d.c generator.
4. The e.m.f of an a.c generator can be converted to different voltages with the help of transformers. This is not possible for e.m.f of a d.c generator because transformers work with a.c only.

The transformer

The transformer is a device that uses mutual induction between two coils to change an alternating voltage across one coil to a larger or smaller alternating voltage across the other coil. A transformer consists of two coils with different number of loops (Figure 2.38), linked by an iron core so that the magnetic flux from one coil passes through the other coil. When the flux generated by one coil changes (as it does continuously in an a.c power source), the flux passing through the other coil will change, inducing a voltage in the second coil.

In a standard transformer, the two coils are usually wrapped around the same iron core, ensuring that the magnetic flux is the same through both coils. The coil that provides the flux is connected to the a.c power source, and is known as the *primary coil.* The coil in which the voltage is induced is known as the *secondary coil.*

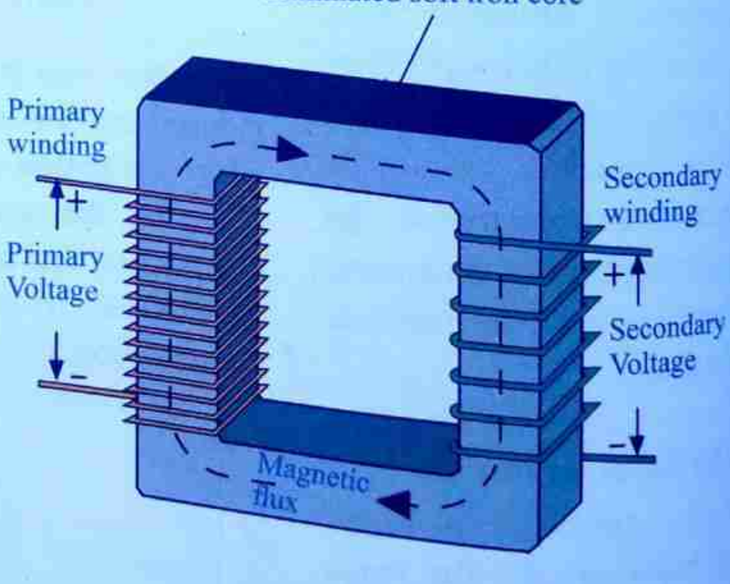


Figure 2.38: *Transformer*

When the number of turns in the secondary coil, N_S is higher than the number of turns in the primary coil, N_P as shown in Figure 2.39, the secondary voltage will be larger than the primary voltage. The transformer whose secondary voltage is larger than the primary voltage is called a *step-up transformer*.

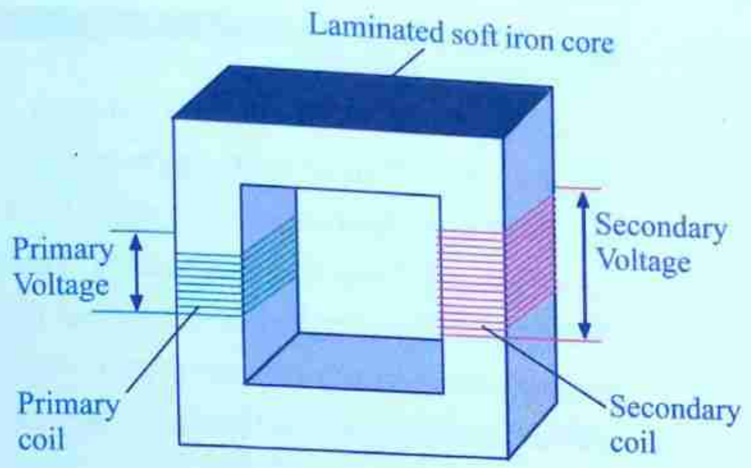


Figure 2.39: *Step-up transformer*

When the number of turns in the secondary coil is less than those in the primary coil as shown in Figure 2.40, the secondary voltage is smaller than the primary voltage. A transformer whose secondary voltage is smaller than the primary voltage is called a *step-down transformer*.

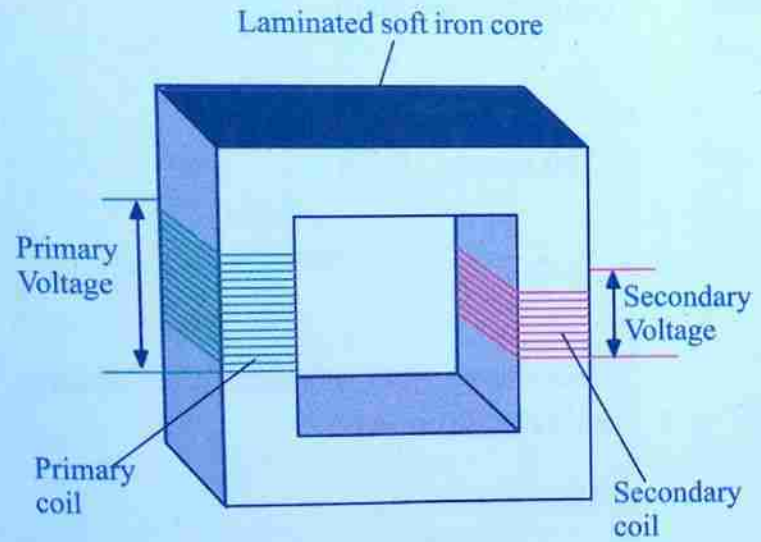


Figure 2.40: *Step-down transformer*

The transformer equation

The e.m.f induced in a coil depends on the number of turns in the coil. It can be shown that the e.m.f is directly proportional to the number of turns making a coil. In an ideal transformer, all the magnetic flux in the primary coil links the secondary coil. The primary e.m.f, V_P is connected to the primary coil of (N_P) turns while the secondary e.m.f, V_S is taken from the secondary coil of (N_s) turns.

Therefore, $V_P \boxtimes N_P$ and $V_S \boxtimes N_S$

This means that, $\dfrac{V_P}{V_S} = \dfrac{N_P}{N_S}$

Assuming that the loss of power in the transformer is negligible, the power in the primary coil, P_P is equal to the power in the secondary coil, P_S. That is,

$$P_P = P_S$$

But $P_P = I_P V_P$ and $P_S = I_S V_S$.

Thus,

$$I_P V_P = I_S V_S$$

where, I_P and I_S are the primary and secondary currents, respectively. Therefore,

$$\frac{I_S}{I_P} = \frac{V_P}{V_S} = \frac{N_P}{N_S} \text{ or } \frac{I_S}{I_P} = \frac{N_P}{N_S}$$

This expression shows that, when the voltage in a transformer is stepped up, the current is stepped down and vice versa.

Transformer efficiency

Transformer efficiency is the ratio of power in the secondary coils P_S to power in the primary coils P_p. The transformer efficiency is normally expressed in percentage.

That is,

$$Efficiency = \frac{P_S}{P_P} \times 100\%$$

But, $P_s = I_s V_s$ and $P_p = I_p V_p$

Thus, $Efficiency = \frac{I_S V_S}{I_P V_P} \times 100\%$

Example 2.1

A transformer is used to step down 240 V mains supply to 12 V for laboratory use. If the primary coil has 600 turns, determine the number of turns in the secondary coil.

Solution

$V_P = 240$ V $\quad V_S = 12$ V

$N_P = 600$ turns $\quad N_S = ?$

$$\frac{N_P}{N_S} = \frac{V_P}{V_S}$$

$$\frac{600 \text{ turns}}{N_S} = \frac{240 \text{ V}}{12 \text{ V}}$$

$$N_S = \frac{12 \text{ V}}{240 \text{ V}} \times 600 \text{ turns}$$

$$= 30 \text{ turns}$$

Therefore, the number of turns in secondary coil is 30.

Example 2.2

A current of 0.6A is passed through a step-up transformer with a primary coil of 200 turns. An output current of 0.1A is obtained from the secondary coil. Determine the number of turns and the voltage across the secondary coil, if the primary coil is connected to a 240 V source.

Solution

$I_P = 0.6$ A $\quad I_S = 0.1$ A

$N_P = 200 \quad V_P = 240$ V

$$\frac{N_S}{N_P} = \frac{V_S}{V_P} = \frac{I_P}{I_S}$$

Using the segment, $\frac{N_S}{N_P} = \frac{I_P}{I_S}$

$$\frac{N_S}{200 \text{ turns}} = \frac{0.6\text{A}}{0.1\text{A}}$$

$$N_S = \frac{0.6 \text{ A}}{0.1 \text{ A}} \times 200 \text{ turns} = 1\,200 \text{ turns}$$

Thus the number of turns in secondary coil is 1 200 turns.

Also,

$$\frac{V_P}{V_s} = \frac{N_P}{N_s}$$

$$\frac{240 \text{ V}}{V_s} \times \frac{200 \text{ turns}}{1\,200 \text{ turns}}$$

$$V_S = \frac{240}{200} \times 1\,200 \text{ V} = 1\,440 \text{ V}$$

Therefore, the magnitude of the secondary voltage is 1 440 V.

Example 2.3

A step-up transformer has 10 000 turns in the secondary coil and 100 turns in the primary coil. A current of 5.0 A flows in the primary circuit when connected to a 12.0 V supply.

(a) Calculate the voltage across the secondary coil.

(b) If the transformer has an efficiency of 90%, what is the current in the secondary coil?

Solution

$I_P = 5\text{ A} \quad I_S = ? \quad N_S = 10\,000\text{ turns}$

$N_P = 100\text{ turns} \quad V_P = 12\text{ V} \quad V_S = ?$

(a) $$\frac{N_S}{N_P} = \frac{V_S}{V_P} = \frac{I_P}{I_S}$$

Using the segment, $\frac{N_S}{N_P} = \frac{V_S}{V_P}$

$$V_S = \frac{N_S}{N_P} \times V_P$$

$$V_S = \frac{10\ 000\text{ turns}}{100\text{ turns}} \times 12\text{ V}$$

$$V_S = 1\,200\text{ V}$$

(b) Power in Primary $(P_p) = I_p V_p$

$$P_P = 5.0\text{ A} \times 12.0\text{ V}$$

$$P_P = 60\text{ W}$$

$$Efficiency = \frac{P_S}{P_P} \times 100\%$$

But power in secondary $(P_S) = I_S V_S$

$$Efficiency = \frac{I_S V_S}{P_P} \times 100\%$$

$$I_S = \frac{Efficiency \times P_P}{V_S \times 100\%}$$

$$I_S = \frac{90\% \times 60\text{ W}}{1\,200\text{ V} \times 100\%}$$

$$= 0.045\text{ A}$$

Therefore, the magnitude of the secondary current is 0.045 A.

An ideal transformer should have an efficiency of 100%. Yet in any practical transformer there are power losses which do not allow 100% efficiency. These losses may be caused by the following factors:

(a) Copper windings

There are losses of power in the primary winding in the form of heat due to the resistance of the primary coil.

(b) Eddy currents

Eddy currents induced in the iron core of a transformer are part of the input energy and change into heat. This effect is however reduced by laminating the iron core to reduce the eddy currents.

(c) Hysteresis

Work is done by the input power to alternately magnetize and demagnetize the iron core of a transformer. This process is repeated continuously as the alternating current flows through the primary coil.

(d) Flux linkage

The other loss is due to poor magnetic flux linkage between the primary and secondary coils causing flux leakages.

(e) Resistance of wires

The wires used to make transformer coils have finite resistance. Therefore, as the current flows through the coils, energy is lost in form of heat.

Uses of transformers

Transformers are used in power stations to step up or step down the level of voltage for transmission to the areas of consumption. The stepping up reduces the current so that losses due to resistance in the transmitting wires are reduced. When power reaches the areas of consumption, it is stepped down to the value required for domestic use. Step down transformers are also used in electrical appliances such as radio, TV sets and cellphone chargers.

Activity 2.8

Aim: To construct a simple transformer.

Material: Iron ring of diameter of about 5 cm, Standard Wire Gauge (SWG) 28 insulated copper wire (1.5 m long), a.c source (0 – 12) V, 2 a.c voltmeters (0 – 6) V, connecting wires, switch, 100 Ω resistor

Procedure

1. Wind 15 turns of the insulated copper wire on the iron ring to cover less than half the circumference of the ring. Strip the ends of the wire and leave them hanging.
2. Using another piece of the same wire, wind 30 turns on the other half of the ring and again leave the ends hanging after stripping them.
3. Connect the other components as shown in Figure 2.41. This is a step up transformer.

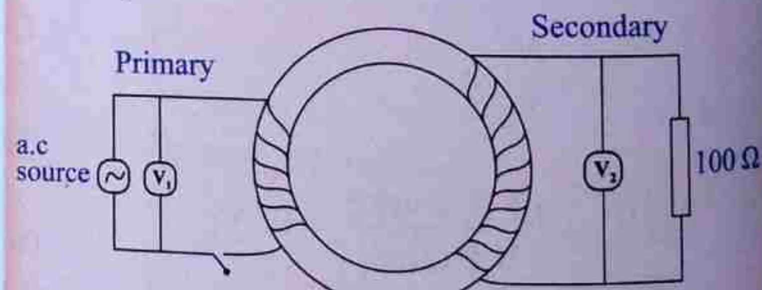


Figure 2.41

4. With the switch open, adjust the a.c. source until the voltmeter V_1 reads 3.0 V. Close the switch and read both voltmeters.
5. Now reverse the coils so that the primary coil becomes the secondary coil. With the switch open, adjust the a.c source until V_1 reads

6.0 V. Close the switch and read both voltmeters.

Questions

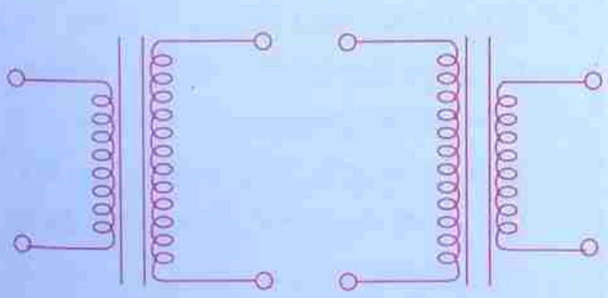
(a) What were the readings on the voltmeters when the switch was closed in step 4?

(b) What were the readings on the voltmeters when the switch was closed in step 5?

(c) Compare the readings in (a) and (b) above with the expected readings if your transformers were ideal. Explain your findings.

(d) Determine the efficiency of your transformers.

When the primary coil has 15 turns and the secondary coil has 30 turns, the input of 3.0 volts should be stepped up to 6.0 volts. When the connections are reversed, the input of 6.0 volts should be stepped down to 3.0 volts. However, outputs will vary depending on the efficiency of the transformer. In circuit diagrams, transformers are represented by the symbol shown in Figure 2.42 (a) and (b).



(a) Step-up transformer (b) Step-down transformer

Figure 2.42: *Transformer symbol*

Exercise 2.3

1. (a) A step-down transformer has 400 turns in the primary coil and 20 turns in the secondary coil. If the primary voltage is 400 V, find the voltage output at the secondary coil.

   (b) When the primary circuit in the transformer in part (a) above has a current of 0.3 A, the secondary circuit is found to have a current of 5.8 A. Determine the efficiency of the transformer.

2. The ratio of the number of turns in the secondary coil in a transformer to that in the primary coil is 16:1. If the current in the secondary circuit is 4.0 A, what is the current in the primary circuit?

3. A 20-watt lamp with a resistance of 7.2 Ω uses a power supply from the secondary coil of a transformer. If the primary coil is connected to a 120 V source,

   (a) what is the current in the lamp when it is switched on?

   (b) what is the secondary voltage?

   (c) what is the ratio of the number of turns on the primary coil to the number of turns on the secondary coil?

   (d) what type of a transformer is it?

Chapter summary

1. Electromagnetism is the phenomenon of interaction of electric field and magnetic field.
2. A current flowing through a conductor generates a magnetic field around the conductor. The strength of the generated magnetic field depends on the amount of current and the dimensions of the conductor.
3. The Right-hand grip rule for a straight conductor states that: "If a wire carrying a current is gripped by the right hand with the thumb pointing in the direction of a conventional current, the fingers will curl around the wire pointing in the direction of the magnetic field".
4. For a solenoid, the Right-hand grip rule states that: "When you wrap your right hand around a solenoid with your fingers pointing in the direction of the conventional current, your thumb points in the direction of the magnetic North pole."
5. Maxwell's cork screw rule states: "If a right-handed screw is rotated such that, its tip advances in the direction of the current, then the direction of screw rotation represents the direction of the magnetic field due to the conductor."
6. A magnetic field exerts a force on a current-carrying conductor. The force on a straight conductor depends on the amount of current, the strength of the magnetic field and the length of the wire.
7. The direction of force exerted by a magnetic field on a current carrying conductor is determined by using Fleming's left-hand rule. The rule states that: "If you hold the index finger, the middle finger and the thumb of your left hand mutually perpendicular to each other so that the index finger points in the direction of the magnetic field and the middle finger points in the direction of current in the conductor, then the thumb will point in the direction of the force acting on the conductor".
8. Two straight parallel conductors repel each other if they carry current in opposite directions and attract each other if their currents are in the same direction.
9. Electromagnetic induction refers to the production of an e.m.f whenever there is a change in magnetic flux.
10. Whenever there is a change in magnetic flux linked with a conductor (coil) the magnitude of induced e.m.f in it is directly proportional to the rate of change of the magnetic flux linking the conductor.
11. The direction of induced e.m.f is determined using Lenz's law of electromagnetic induction. The

law states that: "The direction of the induced e.m.f is such that the resulting induced current flows in a direction that opposes the change that causes it".

12. Self-induction is the production of an e.m.f in a conductor as a result of changing (increasing or decreasing) current in the same conductor.

13. Mutual induction is the production of an e.m.f in one conductor as a result of a changing (increasing or decreasing) current in another conductor, placed close to the first conductor.

14. A generator is a device that uses a coil of wire rotating in an external magnetic field to produce either an alternating current or a direct current.

Revision exercise 2

Choose the most correct answer in items 1 to 4.

1. Figure 2.43 shows a coil placed between the poles of a magnet.

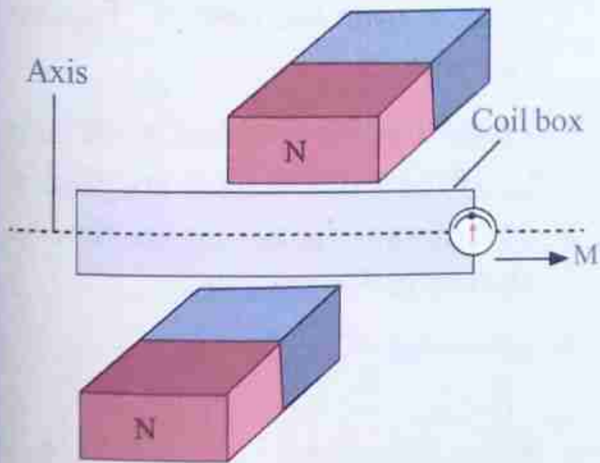


Figure 2.43

The coil is moved between the magnets as follows:

(a) Back and forth in the direction of a vector M.

(b) Towards N-pole and away from N-pole.

(c) Up and down.

(d) Rotated about the axis shown by the dotted line.

In which two movements does the galvanometer show deflection?

(i) (a) and (b) (ii) (a) and (c)

(iii) (b) and (c) (iv) (b) and (d)

2. Which of the following is true about generators?

(a) A.c generators use electromagnets and d.c generators use permanent magnets.

(b) A.c generators have slip rings and d.c generators have commutators.

(c) In a.c generators, the coil is wound on a laminated soft-iron core and in d.c generators on a solid core.

(d) In a.c generators the field coil moves and in d.c generators the armature moves.

3. What is the purpose of the capacitor in an induction coil?
   (a) To store charge.
   (b) Reduce pressure at contact points.
   (c) Create a p.d across contact points.
   (d) Produce a spark at the contact points.
4. Which of the following is a correct method for inducing e.m.f in a wire?
   (a) Moving the wire between the poles of a magnet parallel to the magnetic field.
   (b) Moving the wire between the poles of a magnet perpendicular to the magnetic field.
   (c) Moving the wire near a battery.
   (d) Moving the wire away from a battery.
5. A compass is placed above a wire AB as shown in Figure 2.44.

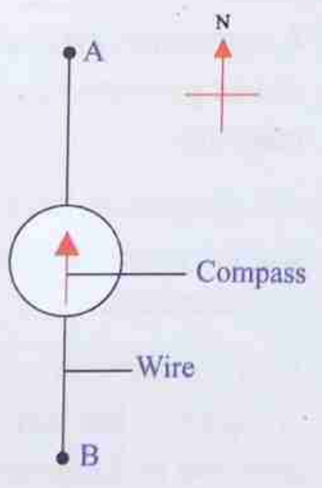


Figure 2.44

   (a) Explain how the compass needle can be made to point due West.
   (b) Explain how the compass needle can be made to point due East.
   (c) Explain how the compass needle can be made to point due South.
6. Two soft-iron rods are suspended inside a solenoid as shown in Figure 2.45.

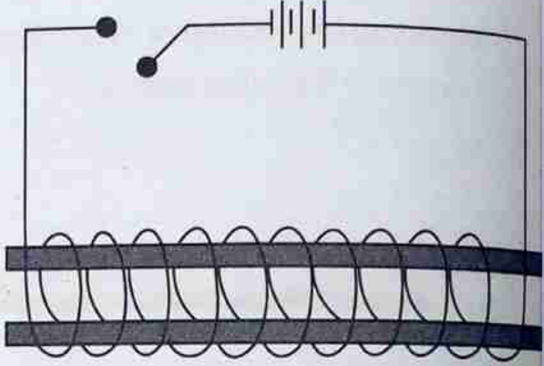


Figure 2.45

When the switch is closed, will the two rods move towards each other or away from each other? Explain your answer.

7. Three parallel wires, A, B and C, of the same length are placed equidistant from one another as shown in Figure 2.46.

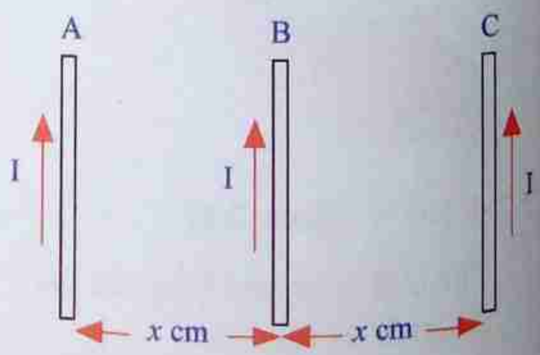


Figure 2.46

Wires A and C are fixed but wire B is free to move. If the same current, I is passed through each of the wires at the same time, in which direction will the wire B tend to move? Explain.

8. Figure 2.47 illustrates a magnet moving relative to a coil of wire and the resulting induced current.

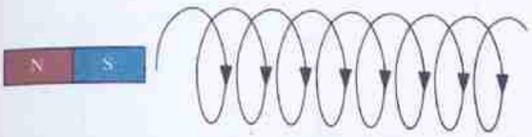


Figure 2.47

Is the magnet moving towards the coil or away from it? Explain your answer

9. Can a transformer be used to increase the voltage of a battery? Explain.

10. What is the ratio of the number of turns required to match an 80 Ω source to a 320 Ω load?

11. A low-voltage outdoor-lighting system uses a transformer to step down the 240 V household voltage to 24 V. The lighting system has 6 lamps with a total resistance of 9.6 Ω.
    (a) What is the current in the primary coil?
    (b) What is the current in the secondary coil?

12. (a) Two transformers T_1 and T_2 are connected as shown in Figure 2.48.

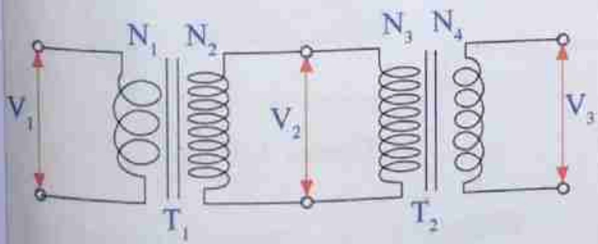


Figure 2.48

Given that $N_1 = 10$, $N_2 = 200$, $N_3 = 100$, $N_4 = 50$, and $V_1 = 240$ V, what is the value of V_3?

(b) An engineer determines that the power input of a transformer is twice the power output. What would be the efficiency of the given transformer?

13. A bar magnet is rotated at a steady rate in front of a wire loop as shown in Figure 2.49.

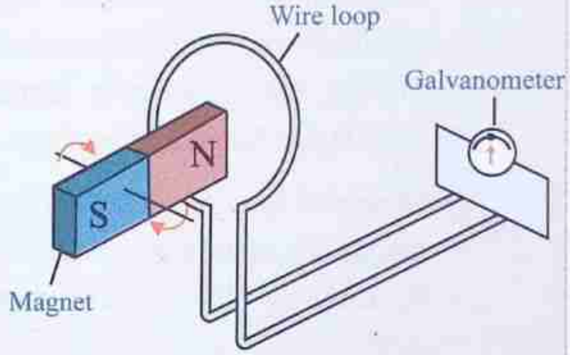


Figure 2.49

Sketch a graph of the induced e.m.f as a function of time at the five different positions of the magnet shown in Figure 2.50.

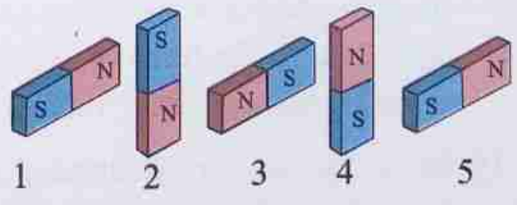


Figure 2.50

14. Figure 2.51 shows a coil in a magnetic field. The coil is rotated in the direction shown by the arrow.

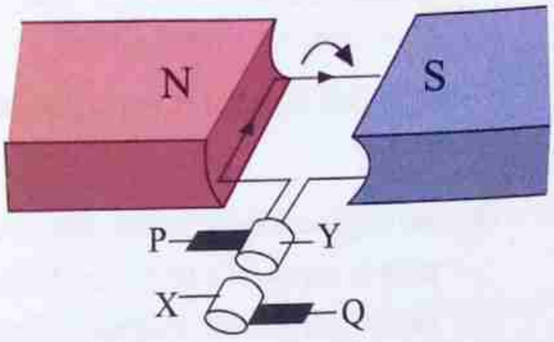


Figure 2.51

(a) Name the parts labelled X and Y.

(b) The terminals P and Q are connected to a cathode ray oscilloscope. Sketch the graph of the electromotive force (e.m.f) produced against time.

(c) State two factors that affect the magnitude of the e.m.f produced by rotating the coil in Figure 2.51.

(d) Explain the changes that should be made in the set-up for it to produce a direct current.

15. The primary coil of a transformer is connected to a 60 V a.c source. The secondary coil is connected to a load of 330 Ω. If the turns ratio is 3:1, calculate the secondary voltage.

16. In a certain transformer, the secondary voltage is one fourth the primary voltage. What is the ratio of secondary current to the primary current?

17. In Figure 2.52, the arrows show the directions of the currents through wires. Copy the wire diagrams and sketch the magnetic field lines around each wire.

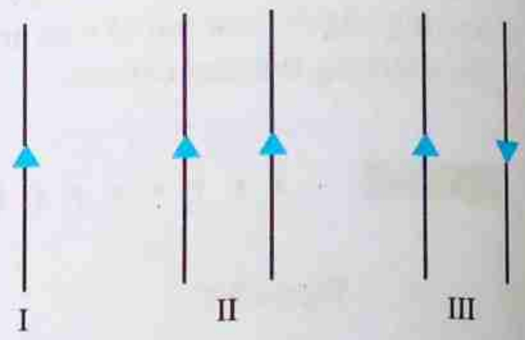


Figure: 2.52

18. In Tanzania, electricity is distributed through out the country by the National Grid System which transmits a.c at a very high voltages. Explain what will happen if:

(a) Direct current is to be used instead of alternating current.

(b) A large current is to be used instead of the high voltage.

19. Figure 2.53 shows a step down transformer connected to a 240 V a.c main supply.

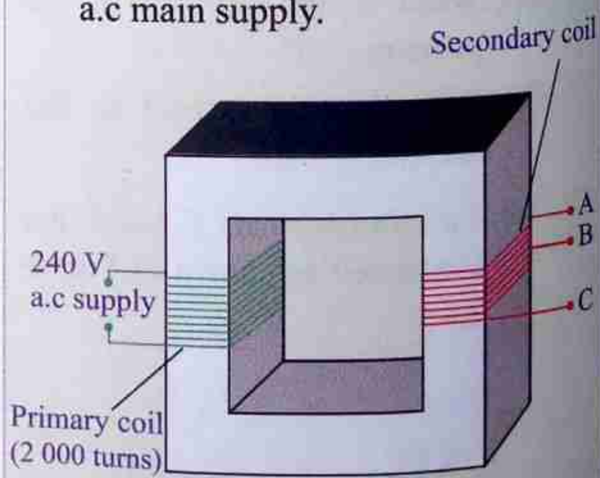


Figure 2.53

When a 6.0 Ω resistor is connected between the terminals A and B, the current flowing in the resistor is 0.4 A, and when the same resistor is connected between terminals B and C, the current is 1.2 A. Assuming the transformer is 100% efficient, calculate:

(a) The p.d between terminals A and B.

(b) The p.d between terminals B and C.

(c) The number of turns on the secondary coil between terminals B and C.

(d) The p.d between the terminals A and C.

(e) The current which would flow in the 6.0 Ω resistor when connected between terminal A and C.

20. Transformers are usually coupled with cooling systems to prevent them from overheating. Explain how heat is generated in a transformer.

Chapter Three

Physics of the atom

Introduction

Everything you see or use in nature is made of matter. Matter is made of atoms which govern the properties of materials. Living creatures keep alive because of oxygen and carbon dioxide gases which are essential in life. Furthermore, hydrogen and oxygen atoms bond to make water, an irreplaceable constituent of life. Generally, our lives depend on atoms. Understanding the structure and behaviour of atoms is very important for humans to be able to manipulate the environment. In this chapter, you will learn about the concept of the atom, natural radioactivity, artificial radioactivity and radiation hazards and safety. You will also learn about nuclear reactions and their applications. The competencies developed will enable you to recognize nuclear radiation and protect yourself from their hazards.

Concept of an atom

The knowledge of the atom dates back to about 25 centuries ago when Greek philosophers started to think about the structure of matter. They proposed that gradual subdivision of matter would result to a smallest indivisible particle. They called this particle "atom", courtesy of a Greek word "atomos" which means indivisible.

However, the properties of atoms remained unknown until the 19th century when Dalton suggested that different elements are made of different atoms and that atoms determine the properties of elements. Yet the structure of the atom remained unknown. All that changed towards the end of the 19th century when scientists strived to understand radioactivity and the nature of light. Several theories were developed to explain the structure of the atom. These include the Thompson's "plum pudding" model of the atom and the Rutherford's planetary model of the atom. Although these models showed that atoms are made of smaller constituents namely, sub-atomic particles, the atom remain to be the smallest particle that can naturally exist. Therefore, one can say, an atom is the smallest particle of an element that dictates the chemical and physical properties of the element and how the element interacts with other elements.

The structure of an atom

The first model of the atom was suggested by J.J Thomson in the 19th century, soon after the discovery of the electron. In his model, Thomson proposed that an atom comprises of a number of negative charges enclosed in a sphere of uniform positive charge. Thomson's structure of the atom resembled a then famous English dessert recipe known as plum pudding. Therefore, the model was named the *plum pudding model.* In this model electrons are embedded in a regular pattern as in a plum pudding and that if the electron is disturbed it retains its original position. The model explained why atoms are neutral. This model also showed that atoms are made of even smaller particles and that the atom is divisible. However, Thomson's model failed to explain some emissions that were observed from atoms. Figure 3.1 illustrates the Thomson's plum pudding model.

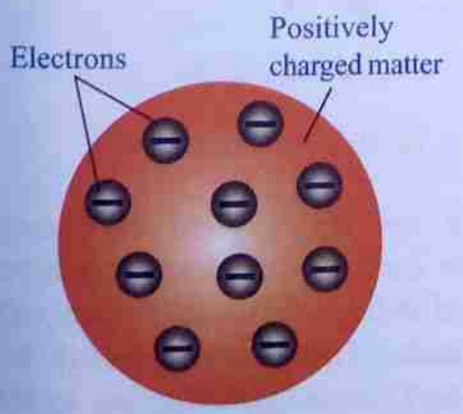


Figure 3.1: *Thomson's model of the atom*

Rutherford's atomic model

Thomson's plum pudding model opened doors for further research on the atomic structure. Rutherford, who was Thomson's student wanted to extend the idea of plum pudding model. He set up an experiment to investigate the structure of the atom by impinging a beam of alpha particles through a gold foil. This experiment was known as the Rutherford's alpha scattering experiment.

Rutherford's alpha scattering experiment

Rutherford and his students (Geiger and Marsden) bombarded a thin sheet of gold foil with a beam of positively charged alpha particles. They placed a circular fluorescent screen around the thin gold foil. The screen was used to investigate the paths taken by the alpha particles after their interaction with the gold foil. Figure 3.2 illustrates the Rutherford's experimental set up.

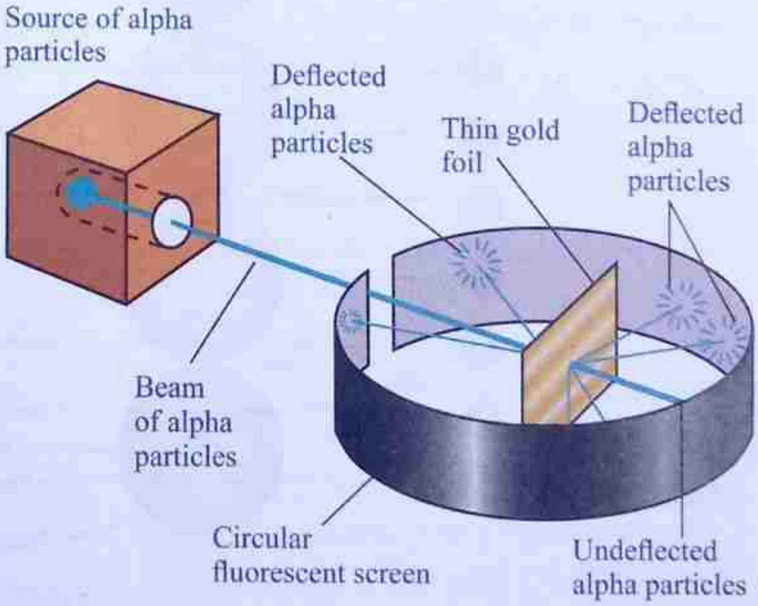


Figure 3.2: *Rutherford's alpha scattering experiment set up*

Rutherford observed that most of the alpha particles passed through the gold foil undisturbed as if the gold foil was an empty space. However, some particles were deflected at large angles. In fact, some particles were deflected backwards as illustrated in Figure 3.3. Such large deflections of alpha particles were not expected on the basis of Thomson's atomic model. This is because, according to Thomson, the positive charge of atoms in the foil is spread such that it is not concentrated enough to deflect alpha particles at large angles. Therefore, Rutherford concluded that the plum pudding model of the atom could not be correct. He then developed a new atomic model based on the following observations from his experiment:

(a) A large fraction of alpha particles that bombarded the gold foil passed through the foil undeflected. This suggested that most part of the atom is an empty space.

(b) Some alpha particles were deflected by the gold foil at very small angles. This suggests that positive charges in an atom are not uniformly distributed. That is, the positive charges are concentrated at the centre of the atom.

(c) Very few alpha particles that passed at the centre of the atom of the gold foil were deflected at an angle that is nearly 180°. This suggests that, the volume occupied by the positively charged particles is very small compared to the total volume of an atom.

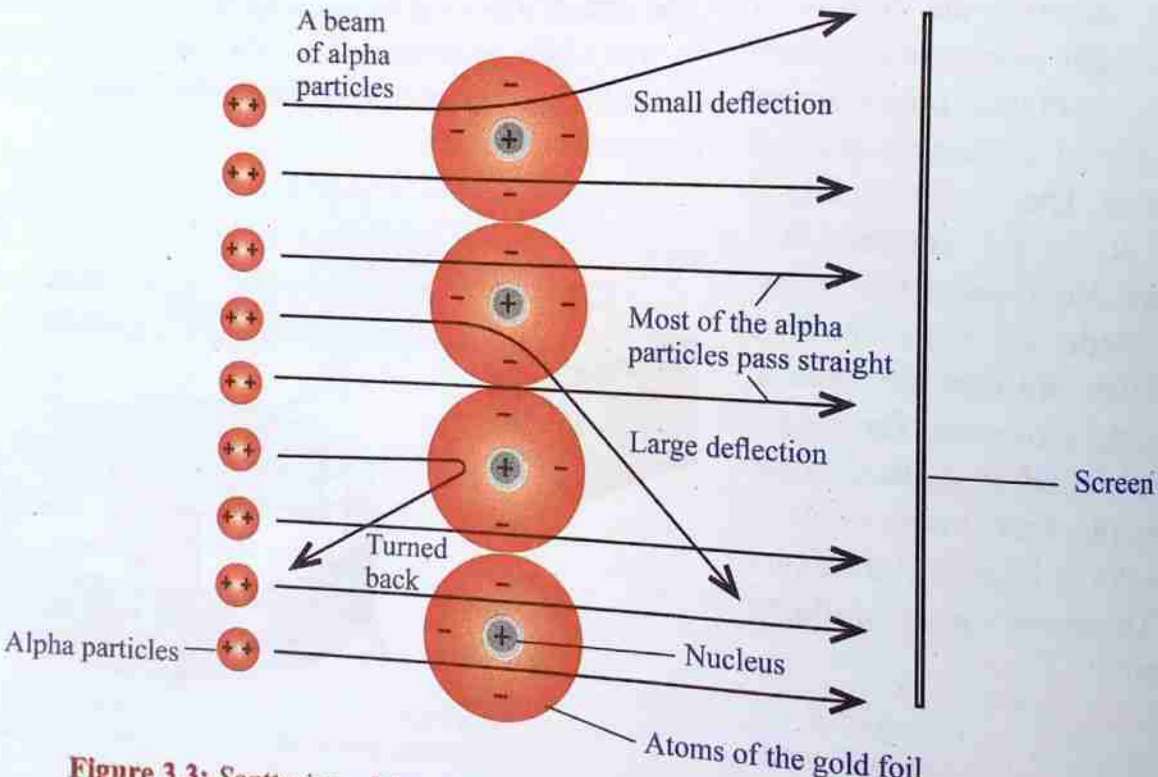


Figure 3.3: *Scattering of the alpha particle according to the Rutherford experiment*

Rutherford's planetary model of the atom

Rutherford used his conclusion to develop an atomic model which is famously known as Rutherford's planetary model. This is because the model borrowed the idea from the arrangement of planets around the Sun. In this model, an atom is made of a tiny, dense and positively charged core. The core is called the *nucleus*. Any electrons belonging to the atom are assumed to be moving in orbits around the nucleus, in the same manner that the planets revolve around the Sun. Electrons are held to the nucleus by electrostatic force of attraction between the nucleus and the electrons. Figure 3.4 illustrates the Rutherford's planetary model of the atom.

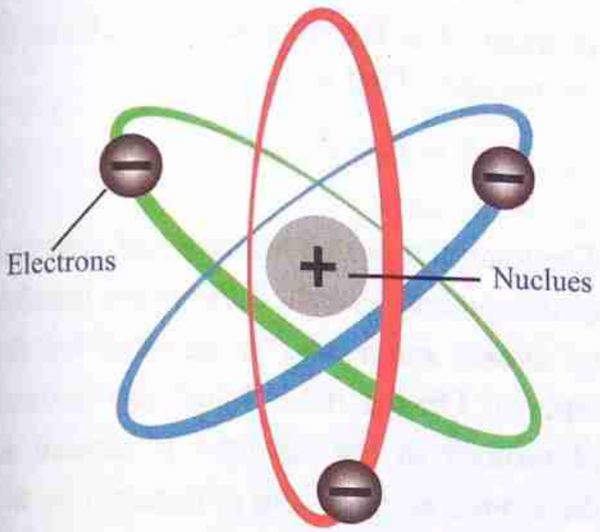


Figure 3.4: *Rutherford atomic model*

Success and limitations of Rutherford's atomic model

The Rutherford's atomic model was very successful in explaining the results of the alpha scattering experiment. The model also explained that the large part of the atom is empty and that the mass of an atom is almost concentrated at the centre of the atom. However, the model has its limitations. These include:

(a) The model could not explain the stability of the atom.

(b) The model was incomplete as it did not describe the arrangement of electrons around the nucleus.

Limitations of the Rutherford's planetary model of the atom were overcome by the Bohr's model of the atom. However, this model is out of scope of this book, so it will not be discussed.

Task 3.1

Discuss the similarities and differences between the Thomson's model and the Rutherford model of an atom. Present your work in class.

Exercise 3.1

1. Basing on Thomson's model of an atom, explain why the atom is neutral.
2. At different times scientists have proposed various descriptions or models of the atom to match the experimental evidence available. Briefly describe:
   (a) The plum pudding model as proposed by J.J. Thomson.
   (b) The planetary model as proposed by Rutherford.
3. For each conclusion given below, state its proof from the Rutherford's experiment.
   (a) Most of the mass of the atom is concentrated in a tiny region called the nucleus.
   (b) The nucleus is positively charged.
4. To find out more about the structure of the atom, Rutherford decided to bombard a thin gold foil with alpha particles. The whole apparatus was in vacuum. Explain the role of the following components of the experimental set up.
   (a) Fluorescent coating on a circular screen
   (b) Circular shape of screen
   (c) Evacuated chamber

The structure of the nucleus

Rutherford's planetary model of the atom shows that a nucleus is at the centre of the atom. The nucleus is made up of tiny particles that are called protons and neutrons. These particles are collectively named as nucleons. The proton is positively charged while the neutron is electrically neutral. That means the overall charge of the nucleus is positive. The mass of a neutron is nearly equal to that of a proton but both particles are approximately 1840 times heavier than an electron. The number of protons in the nucleus is the same as the number of electrons revolving the nucleus, making the atom electrically neutral since the charges cancel out. Figure 3.5 describes the structure of the nucleus.

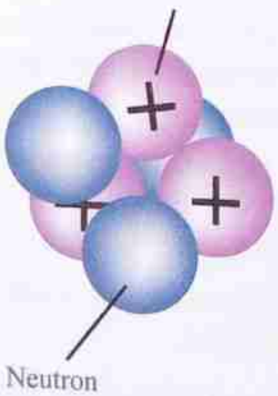


Figure 3.5: *Structure of the nucleus*

Nuclear Notation

Standard nuclear notation involves a chemical symbol, the number of protons and the number of neutrons. The number of electrons is not included in the nuclear notation. The number of neutrons is denoted by N, the number of protons is denoted by Z as a subscript and the number of nucleons in the nucleus is denoted by A as a superscript. For example, if an element symbolized as X, has Z protons and A nucleons, then it is denoted as, ${}^{A}_{Z}X$. Here X is a chemical symbol of the element, A is the number of nucleons in the nucleus. That is,

$$A = Z + N$$

The total number of protons and neutrons in the nucleus is called the *mass number* or *atomic mass* and is denoted by the letter A. On the other hand, the number of protons in the nucleus is termed as the *atomic number* and is denoted by the letter Z.

For example, consider the element Carbon-14. From the periodic table it is found that carbon has atomic mass 14 and atomic number 6. Therefore, the nuclear notation of Carbon-14 is ${}^{14}_{6}C$. The number of neutrons in the Carbon-14 nucleus is given by:

$$N = A - Z$$
$$= 14 - 6$$
$$N = 8$$

Thus, Carbon-14 has 8 neutrons.

The atomic number, Z determines the chemical properties of the element while the mass number, A determines the physical properties of the element. For example, the ability of an element to chemically react with other elements is governed by the atomic number.

On the other hand, the volume of an element depends on its mass number. Scientists arranged all elements according to their chemical properties in a very important tool known as periodic table. In this table elements are indicated using their chemical symbols and atomic numbers.

Note that, protons and electrons are respectively positively and negatively charged particles. A neutron is an electrically neutral particle. The mass of proton, m_p is 1.673×10^{-27} kg and the mass of a neutron, m_n is 1.675×10^{-27} kg. Moreover, the mass of an electron, m_e is 9.109×10^{-31} kg. The charge of a proton is 1.602×10^{-19} C while that of the electron is -1.602×10^{-19} C.

Task 3.2

Study the following elements in the periodic table.

(a) Na (b) U (c) Ge

(i) Write their nuclear notation.

(ii) In each case, indicate the value of A, Z and N.

(iii) Calculate the total charge of each nucleus.

Isotopes of an element

As early as 1910, the English scientist *Frederick Soddy*, who had worked with Rutherford, observed that there were certain elements with identical chemical properties, but different physical properties. Such elements are placed in the same position in the periodic table. Soddy called them *Isotopes*, which is a Greek word meaning "occupying the same place".

Thus, isotopes are atoms of the same element having the same atomic number, Z, but different mass number, A.

For instance, the nuclei represented by ${}^{A_1}_{Z_1}Y$ and ${}^{A_2}_{Z_1}Y$, are two isotopes of the element Y. Since, in an atom the neutron number, N is given by $N = A - Z$, it follows that, the isotopes of an element have different neutron numbers, N. The nuclei of atoms of a specific isotope are called nuclides.

For example, hydrogen has three isotopes denoted by the symbols:

$${}^{1}_{1}H,\ {}^{2}_{1}H \text{ and } {}^{3}_{1}H$$

These isotopes are called hydrogen, deuterium and tritium, respectively. The nuclei of these isotopes have zero, one and two neutrons, respectively. However, each nuclide has only one proton. Other examples of isotope are those of oxygen given by,

$${}^{16}_{8}O,\ {}^{17}_{8}O \text{ and } {}^{18}_{8}O$$

These are isotopes of oxygen with eight, nine and ten neutrons respectively in their nuclei, but all having eight protons.

Isotopes are normally referred by their mass numbers. For example, the three isotopes of oxygen are referred as Oxygen -16, Oxygen -17, and Oxygen -18. Similarly, the isotopes of carbon are referred as Carbon -12 and Carbon -14. These are symbolized as,

$^{12}_{6}C$ and $^{14}_{6}C$, respectively.

Example 3.1

One isotope of chlorine has the symbol,

$$^{37}_{17}Cl$$

(i) Determine the number of neutrons in this isotope.

(ii) Calculate the charge, in coulombs, on the nucleus of this isotope.

Solution

In this chlorine isotope the mass number $A = 37$ and atomic number $Z = 17$.

(i) If N is the number of neutrons in the nucleus, then $A = Z + N$, that is;

$$37 = 17 + N$$

Hence, $N = 37 - 17 = 20$

Thus, the number of neutrons in the nucleus of $^{37}_{17}Cl$ is 20.

(ii) The atomic number $Z = 17$ means that the nucleus of $^{37}_{17}Cl$ has 17 protons.

Each proton has

1.602×10^{-19} C. Hence, the total charge on the nucleus of $^{37}_{17}Cl$ is;

$$Q = +Ze = 17 \times 1.60 \times 10^{-19}\text{C}$$

$$= 2.7 \times 10^{-18}\text{C}$$

Thus, the total charge on the nucleus of this isotope is $Q = 2.7 \times 10^{-18}$ C.

Example 3.2

The element tin has a total of twenty-five isotopes with the lightest isotope being $^{108}_{50}Sn$. Assuming that all twenty-five isotopes exist and the difference in mass numbers of two successive isotopes is one, write down the symbol for the heaviest isotope.

Solution

Different isotopes will differ at least by 1 neutron. That is,

$A_1 = 108, A_2 = 109, A_3 = 110 \ldots A_n$

Since there are 25 isotopes, $n = 25$

From arithmetic progression,

$$A_n = A_1 + (n-1)d$$

In this case, $A_1 = 108$, n = 25, $d = 1$

$$\therefore A_n = 108 + (25-1)1 = 132$$

Denote the heaviest isotope by the symbol,

$$^{A}_{Z}Sn$$

Hence, the heaviest isotope of tin will be $^{132}_{50}Sn$.

Exercise 3.2

1. Draw a well labelled diagram showing the atomic structure of Hydrogen.
2. Explain the following terms:
   (a) Mass number of an atom
   (b) Atomic number
3. (a) What is an isotope?
   (b) Mention the three isotopes of hydrogen.

4. If the number of electrons in an atom is 8 and number of protons is also 8,
   (a) what is the atomic number of the atom?
   (b) what is the net positive charge of the atom?
5. Compare the properties of electrons, protons and neutrons.
6. One isotope of sodium has the symbol, ${}^{24}_{11}\mathrm{Na}$
   (a) Calculate the number of electrons and neutrons in this isotope.
   (b) Calculate the charge, in coulombs, of the nucleus of this isotope.

Stable and unstable Nucleus

With the exception of hydrogen, the nuclei of all atoms contain two or more protons. According to the law of charges, like charges repel while unlike charges attract. It follows that protons within the nucleus repel. However, the repulsive force between protons is overcome by a very strong attractive nuclear force between the nucleons. This attractive force is called the *strong force*. The other nuclear force is the *weak force*. The strong force is responsible for holding the nucleons together. It provides the nuclear *binding energy*. The binding energy is important for providing stability of a nucleus. When the binding energy is large enough, the nucleus is stable. Some isotopes have small binding energy and are said to be unstable. In other words, an atom is stable if the forces within the nucleus are balanced. Conversely, an atom is unstable if the nuclear forces do not balance such that the nucleus has excess internal energy.

The binding energy of the nucleus depends on the number of nucleons within the nucleus. Large number of nucleons contributes to instability of the nucleus. On the other hand, instability may result from excess of either neutrons or protons. This unstable nucleus strives to reach stability by ejecting some nucleons or releasing energy.

Other nuclides are unstable and transform to form other nuclides by emitting particles and electromagnetic radiation. This process is called radioactivity and elements which undergo *radioactivity* are called radioactive elements. This process can be natural or artificial.

Radioactivity *is a random process by which an unstable nucleus transforms to a more stable nucleus by emission of nuclear radiation.*

Natural radioactivity

The transformation of the nucleus is sometimes known as the nuclear decay or disintegration. This process occurs naturally, hence named *natural radioactivity*. When the nucleus disintegrates, it does so with the emission of nuclear radiation which may be alpha particles, beta particles or gamma rays. The decaying nucleus is called the parent nucleus while the new nucleus formed after disintegration is called the daughter nucleus. For example, Cobalt-60 (parent) decays to form Nickel-60 (daughter). The decay process can be classified depending on the ratio of the number of protons to that of neutrons in the parent nucleus.

Types of nuclear decay

Depending on the total number of nucleons in the nucleus and the ratio of the number of neutrons to that of protons (N/Z), there are three common types of nuclear (radioactive) decay, namely: alpha decay (α-decay), beta decay (β-decay) and gamma decay (γ-decay).

(i) Alpha decay

When a nucleus has too many nucleons, the binding energy is reduced and the nucleus becomes unstable. The nucleus then seeks stability by reducing the number of nucleons through emission of an α-particle. An α-particle is a helium nucleus (^4_2He) made of two protons and two neutrons.

Thus, when a nucleus decay by emitting an α-particle, it discards four nucleons. That is, its neutrons, N and protons, Z decrease by 2 and so atomic mass, A decreases by 4. If a parent nuclide ^A_ZX decays by alpha emission it forms a daughter nuclide according to the reaction,

$$\underbrace{^{A}_{Z}\text{X}}_{\substack{\text{Parent}\\\text{nuclide}}} \rightarrow \underbrace{^{A-4}_{Z-2}\text{Y}}_{\substack{\text{Daughter}\\\text{nuclide}}} + \underbrace{^{4}_{2}\text{He}}_{\alpha-\text{particle}}$$

Examples

(i) $^{238}_{92}\text{U} \rightarrow {}^{234}_{90}\text{Th} + {}^4_2\text{He}$

(ii) $^{226}_{88}\text{Ra} \rightarrow {}^{222}_{86}\text{Rn} + {}^4_2\text{He}$

Note that, α-decay is possible whenever the mass of the parent nuclide is greater than the total mass of daughter nucleus and the emitted alpha particle. The excess mass is used as the kinetic energy of the emitted α-particle and the daughter nucleus.

(ii) Beta decay

When a nucleus has too many neutrons or too many protons, it is unstable. Such a nucleus decays by emission of β-particles. In a β-decay the nucleus converts a neutron to a proton and emits a β-particle. A β-particle is an electron formed in the nucleus when a neutron transforms into a proton. β-decay is accompanied with the emission of a neutral particle called the antineutrino ($\overline{\upsilon_e}$). Therefore, the atomic number, Z increases by one while number of neutrons, N decreases by one and the mass number, A remains the same. That is, if ^A_ZX is an unstable nucleus with excess neutrons, then;

$$\underbrace{^{A}_{Z}\text{X}}_{\substack{\text{Parent}\\\text{nuclide}}} \rightarrow \underbrace{^{A}_{Z+1}\text{Y}}_{\substack{\text{Daughter}\\\text{nuclide}}} + \underbrace{^{0}_{-1}\text{e}}_{\beta-\text{particle}}$$

Examples of β-decay:

(i) $^{14}_{6}\text{C} \rightarrow {}^{14}_{7}\text{N} + {}^{0}_{-1}\text{e}$

(ii) $^{60}_{27}\text{Co} \rightarrow {}^{60}_{28}\text{Ni} + {}^{0}_{-1}\text{e}$

This type of decay is called the beta-minus decay and occurs whenever the mass of the parent nuclide is larger than that of the daughter nuclide.

(iii) Gamma decay

Like electrons in their atomic orbit, protons and neutrons in the nucleus are arranged in energy levels each with specific energy. When the nucleus has too much energy, it releases the excess energy by emitting γ-rays. γ-rays discard the excess energy without changing any of the particles in the nucleus. Normally, after radioactive decay, the daughter nucleus has excess energy and therefore it undergoes γ-decay. If a nucleus ${}^{A}_{Z}X$ undergoes β-decay, a new nucleus ${}^{A}_{Z+1}Y^{*}$ which is excited is formed. That is;

$$\underbrace{{}^{A}_{Z}X}_{\text{Parent nuclide}} \rightarrow \underbrace{{}^{A}_{Z+1}Y^{*}}_{\text{Daughter with excess energy}} + {}^{0}_{-1}e + \overline{\upsilon}_{e}$$

This reaction is then followed by γ-decay,

$${}^{A}_{Z+1}Y^{*} \rightarrow \underbrace{{}^{A}_{Z+1}Y}_{\text{De-excited element}} + \underbrace{\gamma}_{\text{Gamma rays}}$$

A good example of gamma decay is in the decay of Cobalt-60.

$${}^{60}_{27}Co \rightarrow {}^{60}_{28}Ni^{*} + {}^{0}_{-1}e + \overline{\upsilon}_{e}$$

Then,

$${}^{60}_{28}Ni^{*} \rightarrow {}^{60}_{28}Ni + \gamma.$$

Note: Nucleus which has * is daughter nucleus in excited state.

Task 3.3

In groups, distinguish between α-decay, β-decay and γ-decay. Present your views for discussion.

Properties of nuclear radiation

The decay of unstable nuclei leads to the emission of nuclear radiation. The radiation is mainly of three types; α-radiation, β-radiation and γ-radiation. These radiations behave differently in different situations. The characteristics of different nuclear radiation are hereby discussed.

(i) Penetrating power

α-particles have low penetrating power, hence can be stopped by a thin sheet of paper. β-particles on the other hand are moderately penetrative as they can penetrate a sheet of paper but can be stopped by a thin aluminium sheet. γ-radiation is highly penetrative as it can penetrate air, paper and aluminium. However, γ-rays can be stopped by several centimetres of lead or concrete. Figure 3.6 illustrates the penetration ability of different nuclear radiation.

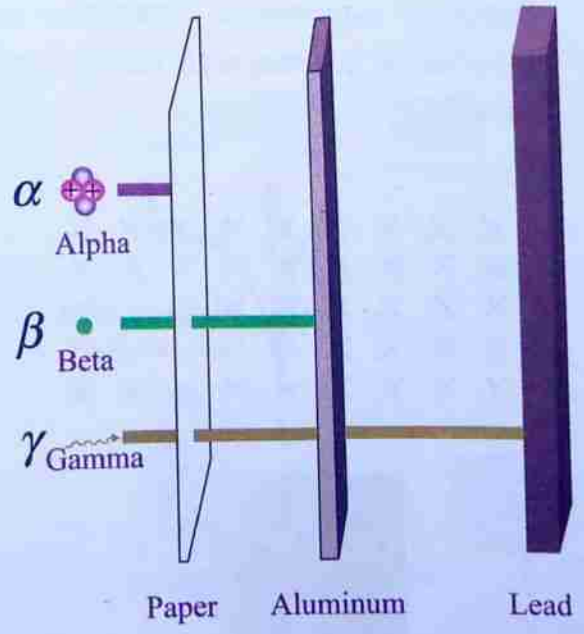


Figure 3.6: *Penetrating power of nuclear radiations*

(ii) Path in an electric field

When passing through a uniform electric field, α-particles are deflected towards the negative plate and β-particles are deflected towards positive plate while γ-radiation pass undeflected. Figure 3.7 shows the path of nuclear radiation in an electric field.

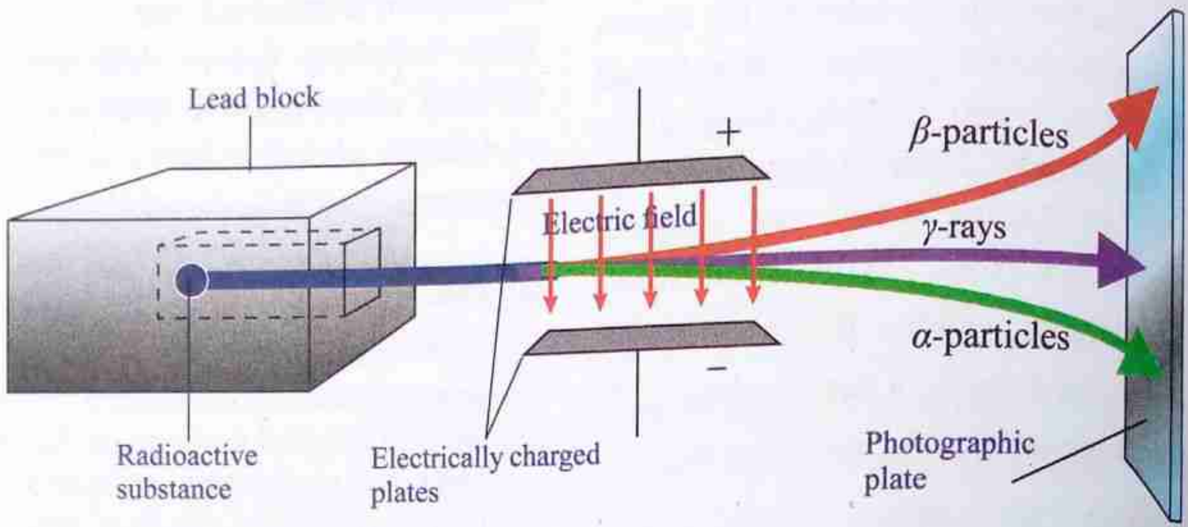


Figure 3.7: *Nuclear radiation paths in an electric field*

(iii) Path in a magnetic field

Upon passing through a uniform magnetic field, nuclear radiation behaves differently depending on the direction of the magnetic field. Figure 3.8 illustrates the respective paths of α-particles, β-particles and γ-radiation in a magnetic field directed into the page.

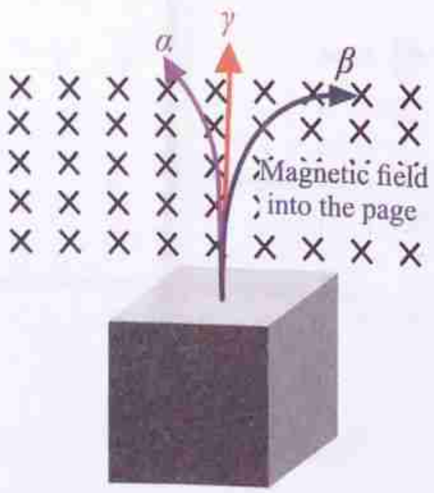


Figure 3.8: *Nuclear radiation paths in a magnetic field*

(iv) Ionising power

α-particles have the highest ionising power among the nuclear radiation types. This is because of their short path in a material. β-particles have a longer path than α-particles in a material leading to a moderate ionising power. γ-rays on the other hand are very penetrative so have the lowest ionising power.

(v) Nature

α-particles are positively charged whereas β-particles are negatively charged. An α-particle is a helium nucleus $\left({}^{4}_{2}\text{He}\right)$ and a β-particle is an electron $\left({}^{0}_{-1}\text{e}\right)$. Therefore, α-particles have a relative charge of $+2e$, β-particles have a relative charge of $-1e$, where e is the electronic charge. γ-radiation consists of photons which are electrically neutral and massless.

(vi) Physical properties

Nuclear radiation types are also characterized by their physical properties. These include mass, relative charge and speed. The masses of α- and β-particles are approximately 6.64×10^{-27} kg and 9.109×10^{-31} kg, respectively. Because of their masses, α-particles and β-particles travel with speeds of 0.1 c and 0.9 c, respectively where c is the speed of light. On the other hand, γ-rays are massless photons which travel with the speed equal to the speed of light, c in vacuum.

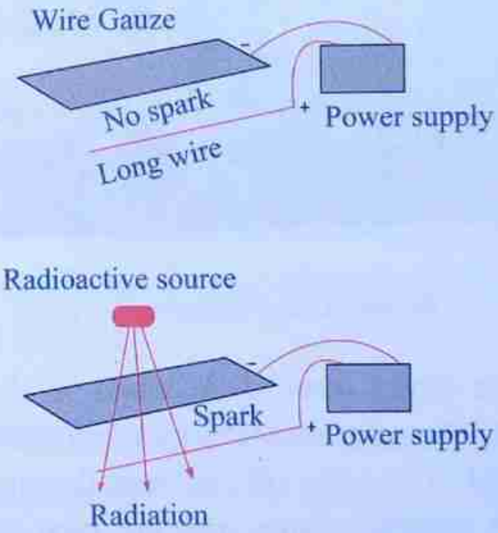
Detection of nuclear radiation

Nuclear radiation cannot be detected by any of the human's five sense organs. An observation of nuclear radiation therefore requires special devices that are named as radiation detectors. A radiation detector is a device that can sense the presence or path of nuclear radiation. Different detectors are designed using different technologies for detecting different types of nuclear radiation. Some detectors such as the spark counter can only detect one or two types of radiation. Other detectors are capable of detecting all radiation types falling on them. Examples of such detectors are the Geiger-Muller counter and the Wilson cloud chamber.

The Spark counter

The spark counter consists of a piece of wire gauze and a long straight wire. The two are separated by a few centimetres of dry air and connected to a high-voltage power supply. The power supply is adjusted to a voltage just below the level required to cause a spark. When a radioactive source is brought nearby, the radiation partially ionises the air between the gauze and the wire. This increases the air's conductivity allowing a spark of electricity between the gauze and the wire (see Figure 3.9 (a)). The sparks produced are recorded by the counter as shown in Figure 3.9 (b).

The spark counter provides a way of visualising the ionising radiation. Its working is based on the ability of nuclear radiation to ionise the air. Since α-particles are the strongest ionising radiation, the spark counter mostly detects alpha radiation.



(a) Production of sparks

(b) Spark counter

Figure 3.9: *Spark counter*

Note: The distance between the source and the wire gauze affects the intensity of α-particles reaching the wire gauze. Again, α-particles have low penetration power, therefore several materials can easily stop the α-particles.

Photographic plate

One of the oldest ways of detecting the ionising radiation is the use of photographic plates. The photographic plate is made of a glass plate coated with a thin transparent material containing silver halide compound (see Figure 3.10).

Figure 3.10: *Photographic plate*

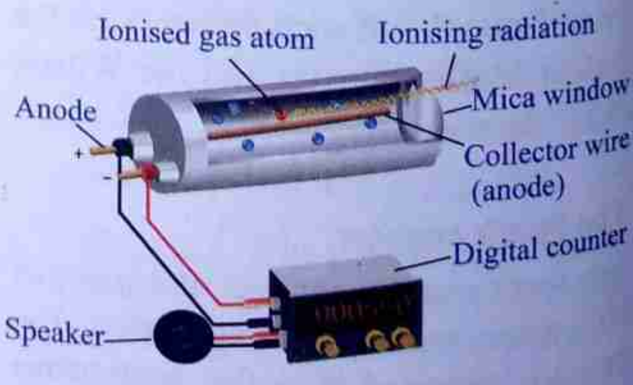
The silver turns dark when it absorbs nuclear radiation as it does when it absorbs light. The level of darkness depends on the exposure time and the type of the radiation it absorbs. The photographic plate is very useful in detecting the presence of ionizing radiation. The plate can only tell the amount of exposure to radiation. The plate can also be upgraded with other technologies so as to measure charge or energy of the radiation. Photographic plates are still used in various applications. For example, the plates are used in X-ray imaging and in making radiation film badges for people working in radiation environments.

However, in some cases, information about the rate of radiation passage is required. In such cases a photographic plate is not useful. Instead, another type of radiation detector that can measure the rate of radiation passage is used. A good example of such a radiation detector is the Geiger-Müller counter.

The Geiger-Müller counter

The Geiger-Müller (GM) counter, illustrated in Figure 3.11(a), is one of the common radiation detectors used for measuring ionising radiation. It is used to detect α-, β- or γ-radiation by taking advantage of their ability to ionise air.

(a)



(b)

Figure 3.11: *The Geiger- Müller counter*

The GM counter consists of a cylindrical tube filled with a noble gas such as helium gas, at low pressure. A metallic wire passes through the centre of the tube and acts as the anode whereas the tube shell serves as the cathode (see Figure 3.11 (b)). One end of the tube is made of a radiation permeable window through which the radiation enters the tube. A high voltage is supplied across the electrodes. When the radiation enters the tube, it ionises air molecules producing both negatively and positively charged ions.

The high p.d between the electrodes accelerates the produced negative charges and positive charges towards the anode and cathode, respectively. This produces a pulse of current in the wire connecting the electrodes. The pulse is counted using a counter or listened using a speaker (see Figure 3.11 (b)). The charged ions are neutralized at the electrodes and the GM tube is ready to count another pulse.

For the GM tube to operate properly, the voltage between its electrodes must be appropriate. Very low voltage will not be able to accelerate the charges while very high voltage will ionise the gas in the tube.

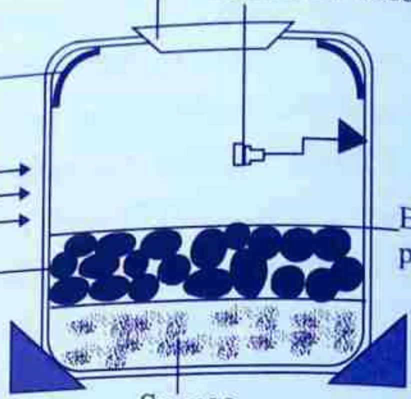
The GM counter is advantageous as it is relatively cheap compared to other counters, durable, portable and can detect all types of nuclear radiation. GM counters are normally used in surveying activities. However, these detectors cannot differentiate which type of radiation is detected. Further, the GM counter has low efficiency and cannot determine the exact energy of the radiation. In order to be able to distinguish between the types of radiation, one needs to use another type of radiation detector named the Wilson cloud chamber.

The Wilson cloud chamber

When radiation passes through highly saturated water vapour in a chamber, it ionises the water vapour. As the particles formed by ionisation pass through the supersaturated water vapour, water droplets tend to condense along the path of the ions. The Wilson cloud chamber takes advantage of this behaviour to detect the path of nuclear radiation across the chamber. Figure 3.12 (a) shows the Wilson cloud chamber.

The Wilson cloud chamber consists of a vessel (chamber) fitted with a piston. This allows dust free air saturated water vapour to be enclosed in the chamber. A camera is placed at an appropriate distance to photograph the series of droplets. The Wilson cloud chamber is therefore a device used to visualise the passage of ionising radiation. Figure 3.12 (b) shows the parts of Wilson cloud chamber.

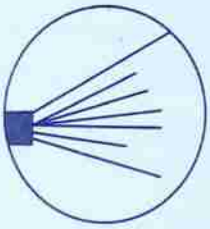
(a)



(b)

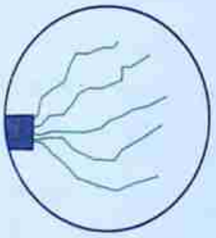
Figure 3.12: *Wilson cloud chamber*

Upon the entrance of ionizing radiation in the chamber, air in the chamber is ionised. The ions produced by the radiation act like condensation centres. The water droplets gather around the ions, thus tracing the path of the ionising radiation. The camera takes photographs of the tracks defined by different shapes depending on the type of radiation that entered the chamber. For example, α-particle's track is broad and straight lines, while that of a β-particle appears to be dotted lines tracks. The γ-rays produce a wave-like track. Figure 3.13 illustrates the tracks produced in the Wilson cloud chamber when different types of nuclear radiation pass through the chamber.



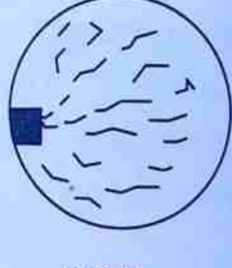
α-particles

Strong ionising power results into straight tracks. The α-particle has a large mass and momentum so it is not easily deflected



β-particles

Very fast beta particle results into thin straight tracks, while the slower beta particles result into short tracks which curve in random direction.



γ-rays

The γ-rays do not produce clear or continuous tracks due to their low ionising power.

Figure 3.13: *Tracks formed by different radiation types in a cloud chamber*

If a cloud chamber is properly used, it can provide information such as the charge, momentum, velocity, lifetime, direction of motion and interaction properties of the incoming radiation. However, one cannot easily deduce the energy of the incoming radiation. The detection of energy of radiation can be achieved by using other radiation detection technologies. These include scintillation detectors and solid state detectors. However, these types of detectors are not discussed as they are beyond the scope of this book.

Half-life of a radioactive nucleus

A radioactive material remains to be radioactive until all the unstable nuclei decay to stable nuclei. Therefore, a radioactive material can be characterised by the type of radiation it emits and the time for which it remains radioactive. The decay of unstable nuclei is probabilistic in nature; therefore, the time that a material remains radioactive is determined by its half-life. At this point, half of the unstable nuclei in a radioactive sample are said to have decayed to stable nuclei. Thus, *half-life is the time taken for the number of unstable nuclei present in a sample at a given time to be exactly halved.* It is normally denoted as, $t_{1/2}$.

Half-life of a given isotope describes how quickly the isotope will decay and for how long it will remain radioactive. Nuclei with long half-lives are less radioactive while those with short half-lives are highly radioactive. For example, Potassium-40, Uranium-238 and Thorium-232 are less radioactive since their half-lives are over one billion years. Other isotopes such as Polonium-214 are highly radioactive. The half-life of Polonium-214 is 0.16 milliseconds.

At any moment, the rate at which unstable nuclei decay is proportional to the number of unstable nuclei. The rate of decay of unstable nuclides is termed as *activity* and denoted by the letter A. This relationship allows for calculations of half-life and hence prediction of how long a sample will remain significantly radioactive.

Suppose the number of unstable nuclei in a radioactive isotope is 1 000 and its half-life is one minute. After one minute, 500 of the nuclei will have decayed. During the next one minute, one half of the remaining 500 nuclei will decay leaving 250. During the third minute, 125 nuclei (one half of 250) will decay and so on. This information can be represented graphically as shown in Figure 3.14.

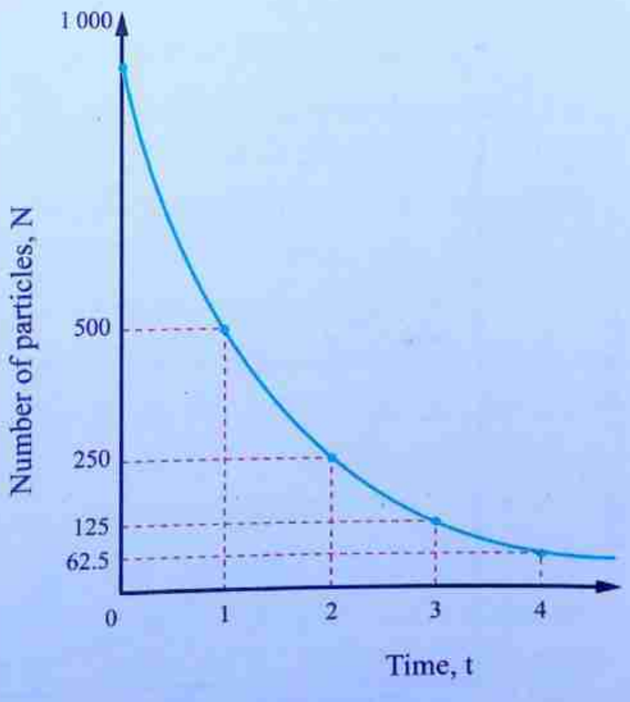


Figure 3.14: *Decay curve of an arbitrary isotope*

Determination of half-life

The number of nuclei disintegrating per second is called the activity of the sample.

That is, $Activity = \dfrac{Number\ of\ disintegrated\ nuclei}{Time\ taken}$

Consider a radioactive sample initially containing N_0 nuclei at $t = 0$ (Figure 3.15).

The amount present after one half-life $\left(t_{\frac{1}{2}}\right)$ is,

$$N_1 = \frac{N_0}{2} = \frac{N_0}{2^1}$$

The amount present after two half-lives $\left(2t_{\frac{1}{2}}\right)$ is: $N_2 = \dfrac{N_0}{4} = \dfrac{N_0}{2^2}$

The amount present after three half-lives $\left(3t_{\frac{1}{2}}\right)$ is: $N_3 = \dfrac{N_0}{8} = \dfrac{N_0}{2^3}$

Therefore, after n half-lives, the amount present will be: $N_n = \dfrac{N_0}{2^n} \Rightarrow \dfrac{N_n}{N_0} = \dfrac{1}{2^n}$

Therefore, $\dfrac{N_n}{N_0} = \left(\dfrac{1}{2}\right)^n$

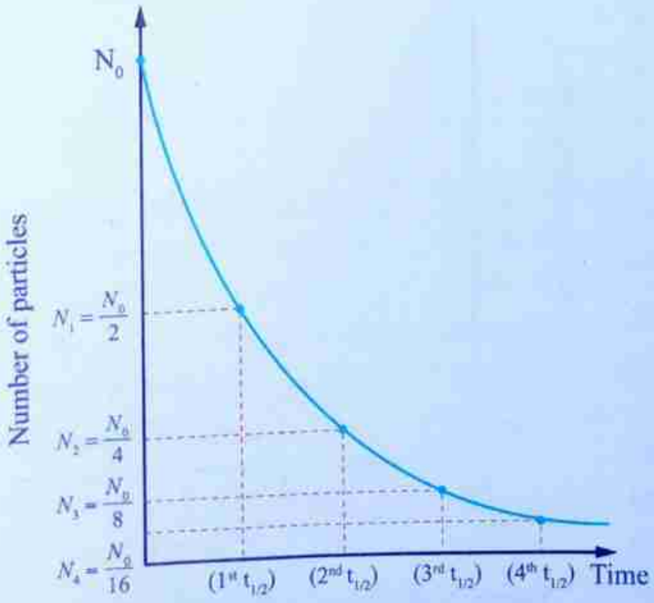


Figure 3.15: *Half-life curve of an arbitrary isotope*

This expression gives a fraction of the nuclei present after n half-lives. If the time required for the nuclei to disintegrate to N_n is t, it can be expressed in terms of half-life as:

$$t = nt_{\frac{1}{2}}$$

$$n = \frac{t}{t_{\frac{1}{2}}}$$

where n is the number of half-lives.

Therefore, $\dfrac{N_n}{N_0} = \left(\dfrac{1}{2}\right)^{\frac{t}{t_{\frac{1}{2}}}}$

Example 3.3

A sample of a radioactive element contains 120 nuclei. How many half-lives must elapse for the sample to get to 15 nuclei?

Solution

$N_0 = 120$ and $N_n = 15$

From $\dfrac{N_n}{N_0} = \left(\dfrac{1}{2}\right)^n \Rightarrow \dfrac{15}{120} = \left(\dfrac{1}{2}\right)^n$

gives, $\left(\dfrac{1}{2}\right)^n = \dfrac{1}{8} = \left(\dfrac{1}{2}\right)^3$

By comparison, $n = 3$.

Therefore, the number of half-lives is 3.

Example 3.4

The half-life of Iodine-131 is 8 days. A sample contains 16 g of Iodine-131. Determine the remaining mass of iodine in the sample after decaying for 24 days.

Solution

$t_{\frac{1}{2}} = 8$ days, $N_0 = 16$, $t = 24$ days.

Using $n = \dfrac{t}{t_{\frac{1}{2}}} \Rightarrow n = \dfrac{24 \text{ days}}{8 \text{ days}} = 3$

Now, from $\dfrac{N_n}{N_0} = \left(\dfrac{1}{2}\right)^n$

$$\Rightarrow \frac{N_n}{16} = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

$$N_n = \frac{16 \text{ g}}{8} = 2 \text{ g}$$

Therefore, the mass of Iodine-131 present in the sample after 24 days is 2 g.

Example 3.5

A sample contains 800 g of Iodine-131. How much Iodine-131 in the sample will remain undecayed after 40 days? (The half-life of Iodine-131 is 8 days.)

Solution

Initial sample, $m_i = 800$ g

Final sample, m_f will be found from:

$$n = \frac{40 \text{ days}}{8 \text{ days}} = 5$$

$$m_f = \frac{m_i}{2^n}$$

$$= \frac{800 \text{ g}}{2^5}$$

$$= \frac{800 \text{ g}}{32}$$

$$= 25 \text{ g}$$

Therefore, 25 g of Iodine-131 will remain undecayed after 40 days.

Exercise 3.3

1. The half-life of ^{60}Co is 5 years. How much time does it take for $\left(\dfrac{3}{4}\right)^{th}$ of its initial mass to disintegrate?
2. How long will it take 600 g of Plutonium-239 (half-life 24 000 years) to decrease to 18.75 g?
3. A certain radioactive element has a half-life of 15.5 hours. If 13.125 g of the element remain undecayed after 62.0 hours, what was the original sample size?
4. Plutonium-239 decays with a half-life of 24 000 years. If Plutonium-239 is stored for 72 000 years, find its remaining fraction.
5. The activity of a radioactive element was measured over a period of 10 minutes using a spark counter. The results in Table 3.1 were obtained:

Table 3.1

Time (s)	Activity (count/s)
0	289
60	243
120	204
180	172
240	144
300	121
360	102
420	86
480	72
540	61
600	51

(a) Plot a graph of activity against time for the sample.
(b) Indicate all the half-lives of the radioactive element on the graph.

Artificial radioactivity

Natural radioactivity occurs spontaneously and randomly, making it unfavorable for many applications of nuclear radiation. This made scientists to come up with the idea of having artificially made radioisotopes. Artificial radioactivity is therefore the radioactivity of isotopes that are artificially produced through bombardment of naturally occurring isotopes with sub-atomic particles. Artificial radioactivity is also called induced radioactivity. The material in which radioactivity is induced is known as the target nucleus while the bombarding material is known as the projectile.

Most radioactive elements are produced artificially by bombarding stable nuclei with high-energy particles. The stable nucleus absorbs energy and becomes unstable. As a result, it decays by emitting nuclear radiation. For example, radioactive Cobalt–60 is produced when Cobalt–59 (the stable and naturally occurring nuclide of cobalt) is bombarded with neutrons.

$$^{59}\text{Co} + {}^{1}_{0}\text{n} \rightarrow {}^{60}\text{Co}$$

Production of artificial radioisotopes

The induced radioactivity is produced in the laboratory by transforming the elements such as Carbon-12 and Phosphorous-15 which are initially stable into their respective unstable isotopes. In turn the unstable isotopes emit radiation, such as alpha, beta and gamma. There are two main methods used to induce radioactivity in materials. These methods are:

(a) Placing a target element in a nuclear reactor, where plenty of neutrons are available. Examples are:

(i) $^{31}_{15}\text{P} + {}^{1}_{0}\text{n} \rightarrow {}^{32}_{15}\text{P}^{*} + \gamma$

(ii) $^{23}_{11}\text{Na} + {}^{1}_{0}\text{n} \rightarrow {}^{24}_{11}\text{Na} + \gamma$

(b) Bombarding the target element with particles from particle accelerators for example,

$$^{23}_{11}\text{Na} + {}^{2}_{1}\text{H} \rightarrow {}^{24}_{11}\text{Na}^{*} + {}^{1}_{1}\text{H}$$

Note: * denotes unstable isotope

Applications of radioactivity

Radioactivity has a wide application in the medical field, food and agriculture, industry, scientific research and many more. The following are some of the applications:

Medical field

In medical field, radioisotopes are used both in diagnosis and treatment of various human diseases. Thus nuclear radiation contributes to the good health and well–being of humans. For example:

(i) Cobalt-60 and Iridium-192 are used in the treatment of cancer. These isotopes emit γ-rays which destroy cancer cells to a great extent.

(ii) Sodium-24 is a radio tracer for diagnostic purposes.

(iii) Iodine-131 is a useful diagnostic tool for assessing thyroid pathologies and treatment of thyroid cancer.

(iv) Iron-59 is used for diagnosis and treatment of anaemia.

Food and Agriculture

In food and agriculture, radioisotopes have made substantial contributions. Uses of nuclear radiation in food and agriculture fields include:

(i) γ-rays from Cobalt-60 are used to destroy microorganisms and parasites in conservation and preservation of food. After being exposed to radiation, certain perishable cereals remain fresh beyond their normal life spans.

(ii) Phosphorus-32 is incorporated in phosphate fertilizer and added to the soil. Phosphorus is taken by the plant for its growth leading to high yield.

(iii) Cobalt-60 is also used in the eradication of insects and pests. Insects and pests are a serious threat to agricultural productivity, as they reduce crop yields and also transmit diseases to the cultivated crops.

Industrial field

Radio-isotopes are commonly used in industry for checking blocked water pipes and cracks as well as detecting leakages in oil pipes. Sodium-24 is placed in a small enclosed ball and allowed to move in a pipe with water. The ball containing radio-isotope is monitored with a detector. If the movement of the ball stops, it indicates a blocked pipe.

Alternatively, Sodium-24 is mixed with oil flowing in underground pipes. With radiation detector, the radioactivity over the pipe is monitored. If there is a leaking point, the radiation detector will show a large activity at that particular point.

Radiation hazards and safety

Radioactive materials have to be handled with care because exposure to radiation can increase the risk of harmful effects to human being. Therefore, safety precautions should be observed to avoid the risks.

Effects of nuclear radiation

A wide variety of reactions occur in response to exposure to nuclear radiation. Some of the reactions occur quickly, while others occur slowly. The killing of cells in affected tissues, for example, may be detectable within minutes after exposure, whereas degenerative changes such as scarring and tissue breakdown may not appear until months or years afterwards.

The damaging effects of radiation on an organ are generally limited to the part of the organ that has been directly exposed.

Consequently, irradiation of only a part of an organ generally causes less impairment in the function of the organ than does irradiation of the whole organ. The following are the effects of exposure to nuclear radiation on different body tissues and organs.

Skin

Radiation can cause various types of injuries to the skin, depending on the dose and conditions of exposure. The earliest outward reaction of the skin is reddening of the exposed area, which may appear within hours after exposure. This reaction typically lasts only a few hours and then, two to four weeks later is followed by one or more waves of deeper and more prolonged reddening in the same area. A larger dose may cause subsequent blistering and ulceration of the skin and loss of hair, followed by abnormal pigmentation a month or years later.

Bone marrow

The blood-forming cells of the bone marrow are among the most radiosensitive cells in the body. If a large percentage of such cells are killed, as can happen when intensive irradiation of the whole body occurs, the normal replacement of circulating blood cells is impaired. As a result, blood cell count may become depressed and infection, hemorrhage or both may occur.

Gastrointestinal tract

The response of the gastrointestinal tract is similar in many ways to that of the skin. Proliferating cells in the mucus membrane that line the tract are easily killed by irradiation, resulting in the denudation and ulceration of the mucus membrane. If a substantial portion of the small intestine is exposed rapidly to a high dose, as may occur in a radiation accident, a fatal dysentery-like reaction results within a very short period of time.

Reproductive organs

Although mature spermatozoa are relatively resistant to radiation, immature sperm-forming cells (spermatogonia) are among the most radiosensitive cells in the body. Rapid exposure of both testes to even a small dose may interrupt sperm production temporarily. A high-dosage exposure may cause permanent sterility in men.

In the human ovary, oocytes (a cell from which an egg or ovum develops by meiosis) of intermediate maturity are more radiosensitive than those of greater or lesser maturity. A low-dose exposure to both ovaries may thus cause only temporary sterility, whereas a high-dosage exposure may cause permanent sterility.

Eyes

Irradiation can cause opacification of the lens, the severity of which increases with the dose. The effect may not become evident, however, until many months after exposure.

Brain and sensory organs

Small doses of radiation can produce phosphene, a light sensation on the dark-adapted retina. The mature brain and nervous system are relatively resistant to radiation injury, but a developing brain is radiosensitive and prone to the damage.

Radiation sickness

The signs and symptoms resulting from intensive irradiation of a large portion of the bone marrow or gastrointestinal tract constitute a clinical feature known as radiation sickness, or acute radiation syndrome.

Early manifestations of this condition typically include loss of appetite, nausea, and vomiting within the first few hours after irradiation, followed by a symptom-free interval that lasts at the main phase of the illness.

The main phase of the intestinal form of the illness typically begins two to three days after irradiation, with abdominal pain, fever, and diarrhoea, which progress rapidly in severity and lead within several days to dehydration, prostration, and a fatal shock-like state.

The main phase of the hematopoietic form of the illness characteristically begins in the second or third week after irradiation, with fever, weakness, infection, and hemorrhage. If damage to the bone marrow is severe, death from overwhelming infection or hemorrhage may follow four to six weeks after exposure unless corrected by transplantation of compatible unirradiated bone marrow cells.

Growth and development of the embryo

Tissues of the embryo, like other tissues, composed of rapidly proliferating cells, are highly radiosensitive. The types and frequencies of radiation effects, however, depend heavily on the stage of development of the embryo or fetus at the time of exposure to radiation. For example, if exposure occurs when an organ is forming, malformation of the organ may result. Exposure during the early stage of embryonic life is more likely to kill the embryo than causing a congenital malformation. However, exposure at a later stage is more likely to produce a functional abnormality in the offspring than a lethal effect or a malformation.

Cancer incidence

Atomic-bomb survivors, certain groups of patients exposed to radiation for medical purposes and some groups of radiation workers have shown dose-dependent increases in the incidence of certain types of cancer. The induced cancers do not appear until years after exposure. However, they do not show distinguishing characteristics by which they can be identified as having resulted from exposure to radiation or from other causes.

Nuclear reactions

When a particle with sufficient energy bombards a nucleus, it may be able to penetrate the nucleus. This may lead to either the formation of a nucleus with greater mass number and atomic number or a decay of the original nucleus. In each case, the re-arrangement of nuclear components occurs. The re-arrangement of nuclear components as the nucleus is bombarded by an energetic particle is called nuclear reaction. There are many types of nuclear reactions and are all subject to conservation laws. The laws involve conservation of charge, momentum, energy and the number of nucleons. Common examples of nuclear reactions are nuclear fission and nuclear fusion.

Nuclear fission

The binding energy of a nucleus gets weaker as the number of nucleons increases. Therefore, nuclei with large number of nucleons are unstable and should decay to more stable nuclei normally by emitting α-particles. However, in some cases emission of α-particles is not conducive and the nucleus just splits to form smaller fragments. The process in which an unstable nucleus of an atom splits into two or more nuclei is called *nuclear fission*. This process may be natural or may be induced by bombarding a heavy nucleus with a neutron. An example of a spontaneous nuclear fission is the fission of Uranium-236.

$$^{236}_{92}U^{*} \rightarrow {}^{92}_{36}Kr + {}^{141}_{56}Ba + 3{}^{1}_{0}n + \text{Energy.}$$

The rate of spontaneous fission is however, very low. For various applications, nuclear fission is induced by bombarding a heavy nucleus with a neutron. An example, of induced nuclear fission is when a Uranium-235 nucleus is bombarded by a neutron as illustrated in Figure 3.16.

$$^{235}_{92}U + {}^{1}_{0}n \rightarrow {}^{236}_{92}U^{*} \rightarrow {}^{144}_{56}Ba + {}^{89}_{36}Kr + 3{}^{1}_{0}n + \text{Energy}$$

$$\text{or } {}^{235}_{92}U + {}^{1}_{0}n \rightarrow {}^{236}_{92}U^{*} \rightarrow {}^{144}_{54}Xe + {}^{90}_{38}Sr + 2{}^{1}_{0}n + \text{Energy}$$

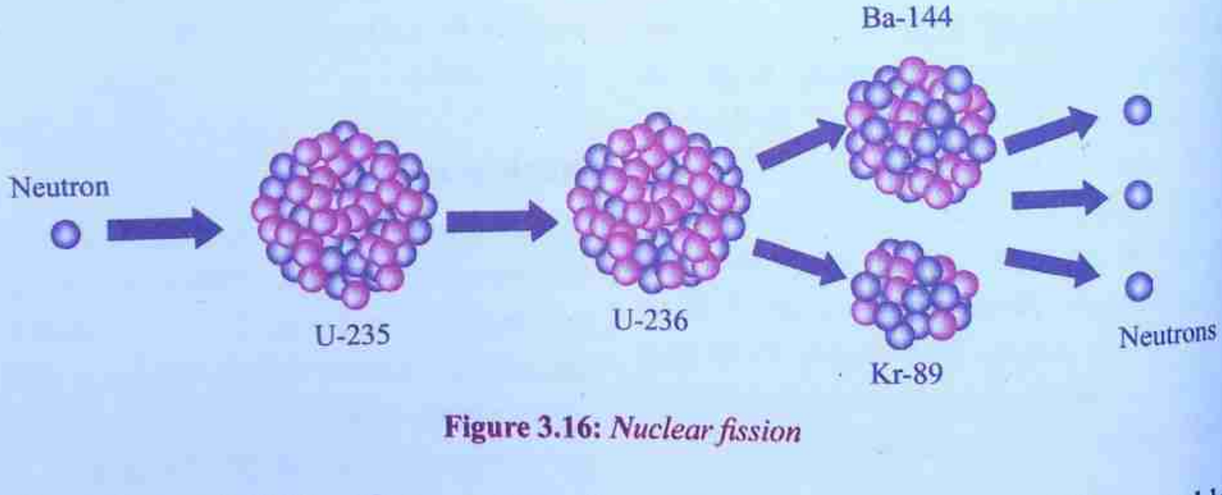


Figure 3.16: *Nuclear fission*

Since, the probability of occurrence of spontaneous fission is very small, the term 'nuclear fission' generally, refers to the induced fission. The products of the nuclear fission are called fission fragments. Although fission fragments have a lower number of nucleons than the parent nucleus, they may still be unstable and hence, continue to decay in different ways. This results to enormous emission of very energetic nuclear radiation. It is for this reason nuclear fission is said to be a dangerous nuclear reaction.

Nuclear chain reaction

When one nucleus is bombarded by a neutron, it releases two or three neutrons. The released neutrons are energetic enough to bombard other nuclei in a material resulting to fission of more nuclei. As time increases, the number of nuclei undergoing fission increases. The reaction that increases in this manner is called *a nuclearchain reaction.* This reaction is useful for producing heat energy in nuclear power plants and nuclear weapons. Figure 3.17 illustrates a nuclear chain reaction.

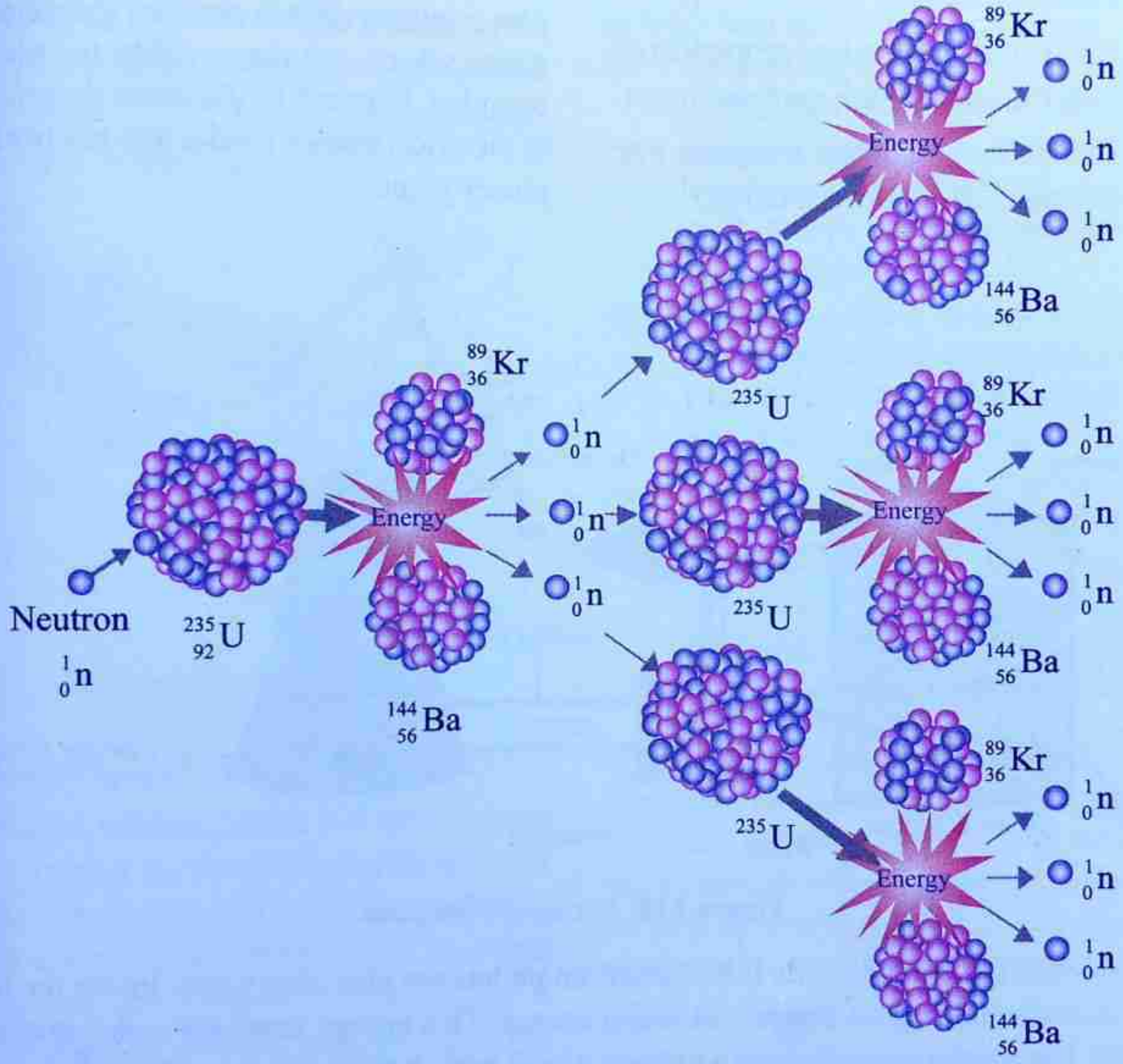


Figure 3.17: *Nuclear chain reaction*

Nuclear fission reaction is highly exothermic. Each nuclear fission releases energy that is approximatly 200 MeV. This energy is in the form of heat and is used mainly for heating water in nuclear power plants.

Applications of nuclear fission

Wide applications of nuclear fission are based on the enormous energy released by the reaction. Some of these applications are:

(a) Propulsion of space crafts.
(b) Generating electricity for industrial and domestic uses.

(c) Sea water desalination.

(d) Production of less common radioisotopes such as Caesium-137 which is useful in radiotherapy.

(e) Power source for propelling submarines and some types of water transport vessels.

(f) Study of structures and properties of materials using high neutron fluxes.

(g) Production of nuclear weapons. For example, the hydrogen bomb.

Note that, nuclear fission provides a cleaner alternative of producing electrical energy. In nuclear power plants, only a small amount of raw material is required to produce a very large amount of energy. This means there is a limited production of wastes. Moreover, no smoke is generated from the nuclear power plant. Thus, nuclear power plants do not produce greenhouse gases which are responsible for global warming. Figure 3.18 illustrates the process of electrical energy production in a nuclear power plant.

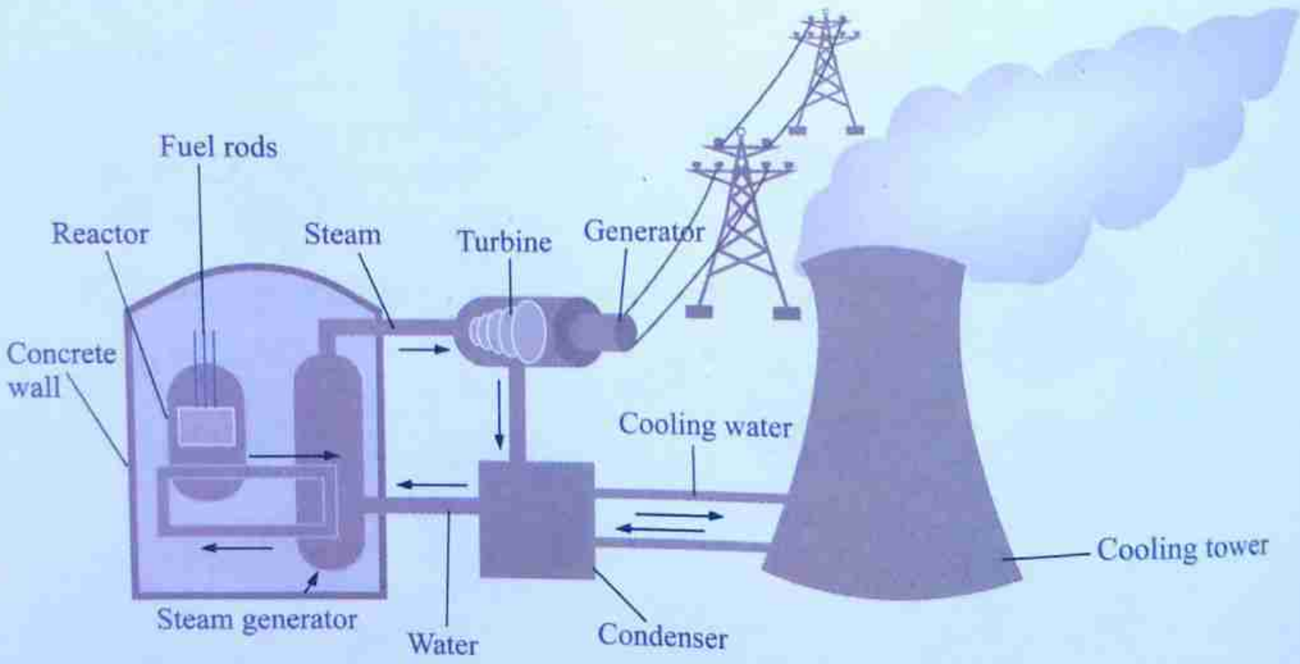


Figure 3.18: *Nuclear power plant*

In a nuclear reactor, fuel rods full of uranium pellets are placed in water. Inside the fuel rod, uranium undergoes fission, releasing energy. This energy heats the water, creating steam. The steam moves through a turbine, which turns a generator to produce electricity. The steam cools back into water, which can then be used over again. At some nuclear power plant, extra heat is released from a cooling tower.

Nuclear fusion

Two small or light nuclei may combine together to form a heavier nucleus. This reaction is called *nuclear fusion*. It is essentially the opposite of nuclear fission reaction. Thus, nuclear fusion is the process whereby light nuclei fuse to form a larger nucleus. A good example, of nuclear fusion reaction is the combination of deuterium and tritium to form helium.

That is,

$$^{2}_{1}H + ^{3}_{1}H \rightarrow ^{4}_{2}He + ^{1}_{0}n + \text{Energy}$$

Unlike fission reaction which is always exothermic, fusion reaction can be exothermic or endothermic. If the product of fusion is a nucleus with mass number less than 60 ($A < 60$), the reaction is exothermic. Stars including the Sun generate energy by nuclear fusion. Figure 3.19 illustrates a nuclear fusion reaction.

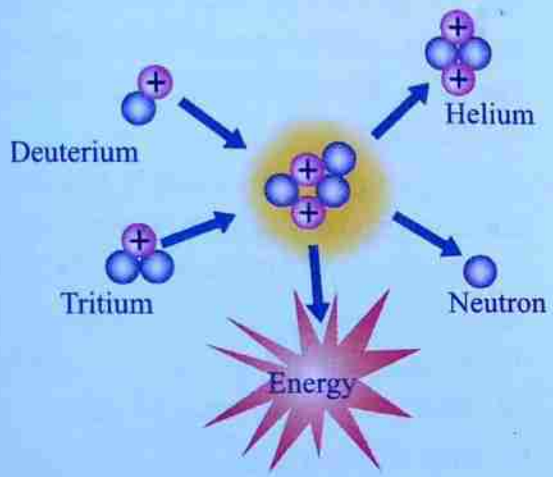


Figure 3.19: *Nuclear fusion*

Applications of nuclear fusion

Artificial nuclear fusion has not yet been developed to the level of using the reaction. This is due to the challenge of maintaining and controlling fusion chain reaction. However, the prospective application is the production of electricity which will be much safer and cleaner than what is offered by nuclear fission reaction. Currently, small scale fusion reaction is used for various research purposes. Note that, nuclear reactors require careful handling because of highly energetic radiation that can cause serious damage. Further, wastes from nuclear reactions are highly radioactive and must be disposed with great caution.

Safety precautions

The guiding principle of radiation safety is "ALARA" which stands for "As Low As Reasonably Achievable". This principle means that even if it is a small radiation dose, if it has no direct benefit, you should try to avoid it. To do this you can use three basic protective measures in radiation safety: time, distance and shielding.

(a) *Time* – simply refers to the amount of time you spend near a radioactive source. If you are in an area where radiation levels are elevated, complete your work as quickly as possible and leave the place.

(b) *Distance* – refers to how close you are to a radioactive source. Maximize your distance from a radioactive source as much as you can.

(c) *Shielding* – to shield yourself from a radiation source, you need to put something between you and the radiation source. The most effective shielding will depend on the kind of radiation coming from the source.

It follows that, protection from the hazards of nuclear radiation includes detecting and reducing exposure. People who work in environments where exposure to radiation cannot be avoided carry devices called *dosimeters* to detect and measure their levels of exposure to radiation.

In order to ensure that radiation protection measures are seriously taken into consideration, special signs are used to warn about the presence of a radiation source.

Radiation warning signs

Radiation warning signs identify radiation hazards through symbol and/or hazard warnings. Warning signs must be posted in each controlled radiation area to indicate the presence of nuclear radiation and the degree of the hazards in the area. Common radiation warning signs are shown in Figure 3.20.

The warning sign shown by Figure 3.20 (a) is a traditional sign used to indicate the presence of a radioactive material or a radiation producing equipment. On the other hand, the sign in Figure 3.20 (b) is relatively new and was introduced to warn people about the presence and hazards of nuclear radiation as well as what should be done. This sign simply means"there is deadly radiation so avoid the area." Both signs are normally posted at the entrance of a room containing a radiation source.

(a) (b)

Figure 3.20: *Radiation warning signs*

Chapter summary

1. An atom is the smallest particle of an element that can naturally exist. It is composed of a nucleus containing protons and neutrons, and surrounded by electrons.
2. The number of protons in the nucleus is called the atomic number and determines the characteristics of the element.
3. The mass number of an atom is equal to the sum of the number of neutrons and the number of protons in a nucleus.
4. Isotopes of an element have the same number of protons in the nucleus but different number of neutrons, that is, the same atomic number but different mass numbers. The isotopes of a given element have the same chemical properties but different physical properties.
5. The nuclei of some isotopes are unstable, making them radioactive. Radioactive nuclei decay by emitting radiation in the form of either alpha particles, beta particles or gamma radiation.
6. Alpha particles are basically helium nuclei composed of two protons and two neutrons and carry a charge of +2e.
7. Beta particles are negatively charged electrons ejected from the nucleus of an atom during radioactive decay.
8. Gamma radiation is a high energy electromagnetic wave emitted from a nucleus.
9. The detection of nuclear radiation is based on its ionising effects. Detection devices include the Geiger-Müller counter, spark counter, photographic film and Wilson cloud chamber.
10. The half-life of a radioactive isotope is the time required for one half of the number of unstable nuclei in a sample to decay.
11. Nuclear radiation has beneficial application in medical, industrial and agricultural fields.
12. Nuclear radiation may cause several health problems if a person is over-exposed to it.

Revision exercise 3

Choose the most correct answer in items 1 to 4.

1. Which of the following statements describes a nuclear fusion reaction?
   (a) Two hydrogen nuclei join to form a helium nucleus.
   (b) A helium nucleus joins with a hydrogen nucleus to form an alpha particle.
   (c) Uranium nuclei split and produce high energy neutrons causing a chain reaction.
   (d) Two helium nuclei join to form a hydrogen nucleus.
2. Which statement is correct about the half-life of a radioactive source?
   (a) It is half the time for the radioactive source to become safe.
   (b) It is half the time it takes an atom to decay.
   (c) It is half the time it takes the activity of the source to decrease to zero.
   (d) It is the time it takes the activity of the source to decrease by half.
3. Which particle is released when $^{30}_{15}P$ decays to $^{30}_{14}Si$?
   (a) Electron
   (b) α – particle
   (c) Neutron
   (d) Positron
4. A radioactive source has a half - life of 80 s. How long will it take for 7/8 of the source to decay.
   (a) 10 s
   (b) 70 s
   (c) 240 s
   (d) 640 s
5. Describe an experiment which led to the discovery of the nucleus of the atom.
6. Explain the significance of the periodic table of elements in relation to radioactivity.
7. One isotope of oxygen has the nuclear notation $^{16}_{8}O$. Determine:
   (a) The number of neutrons in the isotope.
   (b) The charge in coulombs on the nucleus of this isotope.
8. An element has three isotopes, the heaviest of which is represented by the nuclear notation $^{18}_{8}X$, where X is a chemical notation of the element.
   (a) Write down the symbol for the lightest isotope.
   (b) What is the name of the element?
9. Describe the three main particles that make up the atom.
10. Write the nuclide notations of the following isotopes:
    (a) Sulphur–34.
    (b) Silver–107.
    (c) Thorium–230.
11. Write the following using the name-mass number format:
    (a) $^{42}_{20}Ca$
    (b) $^{198}_{80}Hg$
    (c) $^{238}_{92}U$

12. Distinguish between natural and artificial radioactivity.
13. Give two applications of artificial radioactivity.
14. Complete the following decay equations.

(a) ${}^{14}_{6}C \longrightarrow$ ____ $+ {}^{0}_{-1}e + \overline{\upsilon}_e$

(b) ${}^{3}_{1}H \longrightarrow$ ____ $+ {}^{0}_{-1}e + \overline{\upsilon}_e$

(c) ${}^{185}_{79}Po \longrightarrow {}^{181}_{77}Ir +$ ____

(d) ${}^{24}_{11}Na \longrightarrow {}^{24}_{12}Mg +$ ____

(e) ____ $\longrightarrow {}^{204}_{82}Pb + {}^{4}_{2}He$

15. Thorium-232 decay series follows the sequence of radioactive decays ultimately producing an isotope which is stable. Indicate the type of decay represented in each step of the series. The first decay step is done as an example.

(a) ${}^{232}_{90}Th \longrightarrow {}^{228}_{88}Ra + {}^{4}_{2}He$

(b) ${}^{228}_{88}Ra \longrightarrow {}^{228}_{89}Ac$ ____

(c) ${}^{228}_{89}Ac \longrightarrow {}^{228}_{90}Th$ ____

(d) ${}^{228}_{90}Th \longrightarrow {}^{224}_{88}Ra$ ____

(e) ${}^{224}_{88}Ra \longrightarrow {}^{220}_{86}Rn$ ____

(f) ${}^{220}_{86}Rn \longrightarrow {}^{216}_{84}Po$ ____

(g) ${}^{216}_{84}Po \longrightarrow {}^{212}_{82}Pb$ ____

(h) ${}^{212}_{82}Pb \longrightarrow {}^{212}_{83}Bi$ ____

(i) ${}^{212}_{83}Bi \longrightarrow {}^{208}_{81}Tl$ ____

(j) ${}^{208}_{81}Tl \longrightarrow {}^{208}_{82}Pb$ ____

${}^{208}_{82}Pb$ is stable.

16. Write down four health effects caused by nuclear radiation.
17. Why is uranium very important in a nuclear power plant?
18. A sample containing caesium-137 isotope has a mass of 20 mg. If the half-life of caesium-137 is 30 years, what will be the mass of the sample in mg after 90 years?
19. A sample of carbon isotope ${}^{14}_{6}C$, has a half-life of 5700 years. What fraction of ${}^{14}_{6}C$ will remain after 11400 years?
20. (a) What is meant by radioactive decay?
    (b) A certain sample with half-life of 8 days contains 16 g of Iodine-131.
    (i) Write an expression to show the decay process of the sample.
    (ii) Use the expression in (i) to sketch a graph and then estimate the mass of sample which will remain undecayed after 20 days.
21. Describe the use of a Geiger-Müller (GM) tube in detecting nuclear radiation.
22. (a) Explain the behaviour of gamma radiation, alpha and beta particles with respect to electromagnetic fields.
    (b) With the aid of a labeled diagram, describe the penetration abilities of gamma radiation, alpha and beta particles.

Chapter Four

Thermionic emission

Introduction

When you visit a hospital, you may see a radiology department. In this department, medical specialists use imaging technology to diagnose and treat various diseases. Radiographers use X-ray tubes to generate X-rays which are used for diagnosing various health conditions. In physics and engineering laboratories a device named the Cathode Ray Oscilloscope (CRO) plays an important role in designing and developing various scientific instruments. Both X-ray tubes and CROs depend on a phenomenon known as thermionic emission. It follows that, the knowledge of thermionic emission is crucial in our everyday life. In this chapter, you will be introduced to the concept of thermionic emission, as well as the structure and the working principle of CROs. You will also learn about X-rays generation and their applications. The competencies developed will enable you to properly use and maintain the CRO, differentiate the cathode rays from X-rays and protect yourself from X-rays hazard.

Concept of thermionic emission

All materials are composed of atoms. An atom is made of a nucleus that is made of protons and neutrons. The nucleus is surrounded by electrons which revolve around it in different orbits. These electrons are held to the nucleus by the binding energy that arise from the force of attraction between the electrons and the protons. The binding energy allows the electrons to revolve around the nucleus in different orbits with different energy levels. At room temperature, the electrons in the atoms of a material stay in their respective orbits because of the binding energy.

However, if a material like metal is heated, electrons in the surface of the material gain thermal energy. If the thermal energy gained is sufficiently high, the electron becomes more energetic than its respective binding energy. Consequently, that electron is ejected from the surface of the material and becomes a free electron.

If the material is heated further, more electrons from the surface of the material are ejected. The ejected electrons are called *thermions* and the emission phenomenon is called *thermionic emission*. Thermionic emission was first discovered by Thomas Alva Edson.

Thermionic emission is the process in which orbital electrons are emitted from the surface atoms of a material when the material is heated.

Factors affecting thermionic emission

Several factors may influence the frequency of occurrence of thermionic emission. That is, the rate of electron emission from the material depends on the following factors:

(a) *Temperature of the material surface*
The higher the temperature of the surface, the higher the rate of electrons emission from the surface.

(b) *The surface area of the material*
The larger the surface area of the material emitting electrons, the higher the rate of emission of electrons.

(c) *The nature of the material surface*
Each material requires a certain amount of external heat energy to just remove an electron from their surfaces. This energy is called *work function* or *threshold energy*. Materials with low work function require less thermal energy to emit electrons from its surface. Thus, the rate of emission from such materials is high even at low temperatures. On the other hand, when the work function is high, a material requires more thermal energy to release electrons from its surface. Thus, the rate of electron emission from materials with high work function is low even at high temperature.

Cathode rays

A thin filament may be used as the cathode in an evacuated glass tube. When an electric current is passed through the filament, green fluorescent light is observed on the glass just behind the positive terminal. The green fluorescent light is the result of a stream of electrons travelling from the filament to the glass wall. Such a stream of electrons is called cathode rays. The name cathode ray comes from the fact that the stream originates from the filament which is used as the cathode. The rays are also named as the electron beam or e-beams.

Production of cathode ray

Cathode rays are produced in a tube known as Cathode Ray Tube (CRT). CRT is a vacuum tube comprising of three parts: electron gun, deflection system, and the screen or displaying system. The electron gun produces electrons using a wire filament by means of thermionic emission. The deflection system is made of two pairs of metal plates. One pair is called Y-plate and the other pair is called X-plate. At the far end, the tube has a fluorescent screen for observing the electron beam. The tube is vacuated in order to prevent electrons from being scattered within the tube. Electron scattering would distort or dim the image formed on the screen. The structure of CRT is illustrated in Figure 4.1 and the functions of each of its components are explained in Table 4.1.

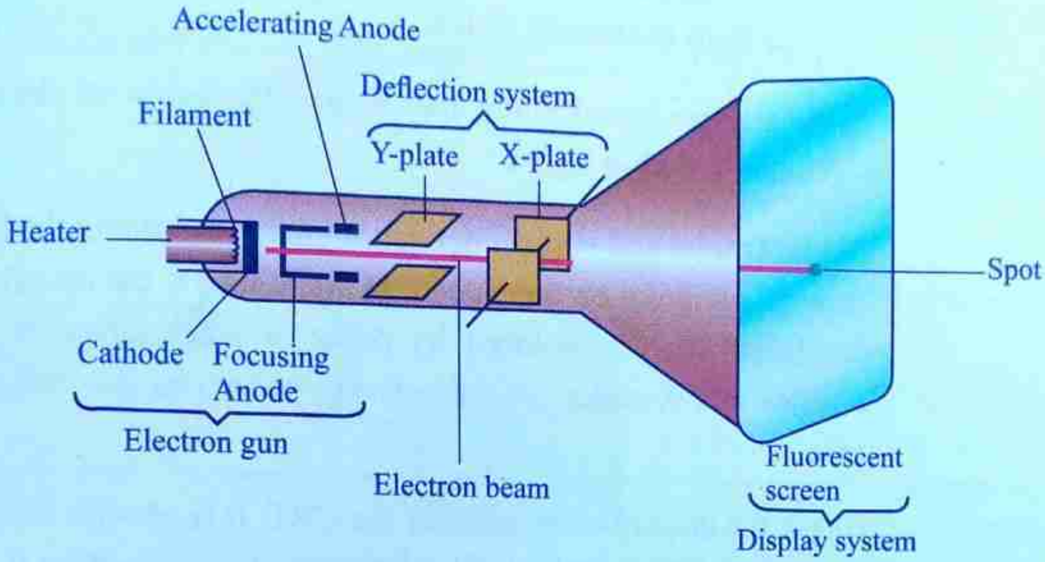


Figure 4.1: *The cathode-ray tube*

Table 4.1: *CRT components and their functions*

Component	Function
Cathode	This is a metal filament such as tungsten that is heated to high temperatures either directly by an electric current or indirectly by a heating element. The cathode can be heated up to several thousand degrees Celsius. At this high temperature, electrons are emitted from the cathode by thermionic emission. Note that the cathode is maintained at a negative voltage.
Focusing anode	This is a metal plate maintained at a high positive voltage. Its main function is to attract electrons and focus them into a beam. It also controls the brightness of the image.
Accelerating anode	This is a metal disk maintained at a high positive voltage (5 000 V to 50 000 V). The anode accelerates the electrons that are ejected from the cathode. There is a small opening in the anode through which a narrow beam of electrons passes and enters a region where their direction can be altered. In some CRTs, there is a negatively charged grid located between the cathode and anode that can limit the number of electrons in the beam. Note that, the cathode and anode are collectively known as the electron gun.

Vertical (Y) deflection plates	These are two parallel metal plates carrying equal but opposite charges. They are used to deflect the electron beam vertically (up or down). The electron beam is attracted to the positive plate and repelled from the negative plate.
Horizontal (X) deflection plates	These are two metal plates that deflect the beam of electrons horizontally (left or right). The horizontal and vertical deflection plates can direct the beam towards any point on the screen. In some CRTs, the electrically charged plates are replaced by poles of electromagnets. The Y-plates and X-plates are collectively known as the deflection system.
Fluorescent screen	This is the display component of the CRT. It is phosphor-coated material so that it emits light whenever electrons strike it. The deflection plates move the electron beam to different positions on the screen resulting in the formation of an image. The image persists for a very short time.

In the cathode ray tube, electrons are produced at the cathode through thermionic emission. The produced electrons are accelerated towards the fluorescent screen by the anode. When the cathode rays hit the fluorescent screen, the screen glows. The horizontal and vertical deflection plates position the beam on the screen.

Task 4.1

Describe how cathode rays are produced. Make a summary and present it to the rest of the class.

Properties of cathode rays

Cathode rays are stream of electrons observed in vacuum tubes. Therefore, the properties of cathode rays are merely, the characteristics exhibited by electrons. The following are the properties of cathode rays:

(a) They travel in straight lines in the absence of electric or magnetic fields.

(b) They carry negative charge.

(c) They cause fluorescence (glow) when they strike a fluorescent material.

(d) They have both energy and momentum.

(e) They are deflected by electric and magnetic fields. The behaviour of cathode rays in an electric and magnetic field are described in Figure 4.2 (a) and (b).

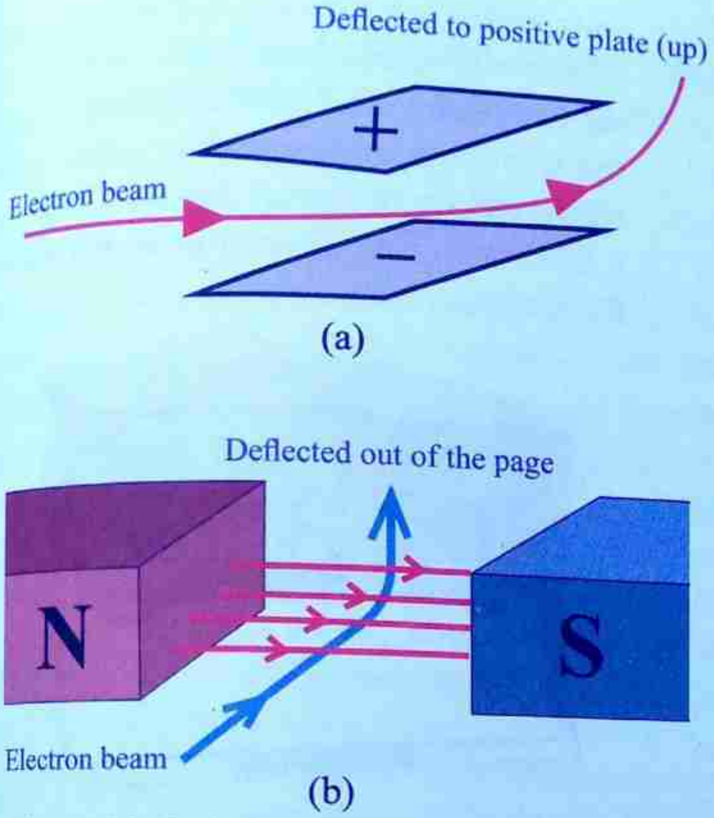


Figure 4.2: *Deflection of cathode rays in (a) an electric field, and (b) a magnetic field*

(f) Cathode rays can ionise gas through which they pass.

(g) These rays can penetrate thin sheets of paper or metal foils depending on their energy.

(h) Cathode rays produce heat upon falling on matter. Thus, they affect photographic plates.

(i) Cathode rays travel with a speed less than that of light in vacuum.

Task 4.2

(a) To broaden your understanding, study the properties and uses of cathode rays. You may use textbooks and internet.

(b) In groups, compile your findings in (a) above and present them on a chart. Present your report to the rest of the class.

Applications of cathode-ray tube

Many devices make use of the CRT. Such devices include laboratory oscilloscope, CRT television, computer and RADAR display. The following are some common devices that use CRT.

The television (TV) and computer display

One important application of the CRT is found in the CRT based TVs. In these televisions, the image is formed on the screen by varying the brightness at thousands of points on the screen. The brightness of a point on the screen depends on the number of electrons that strike it. The intensity of the electron beam can be varied by changing the voltage on the grid located between the cathode and the anode.

The horizontal deflection plates cause the electron beam to scan across the screen from right to left whilst the vertical deflection plates move the beam up or down the screen. Figure 4.3 shows a TV set that uses the CRT.

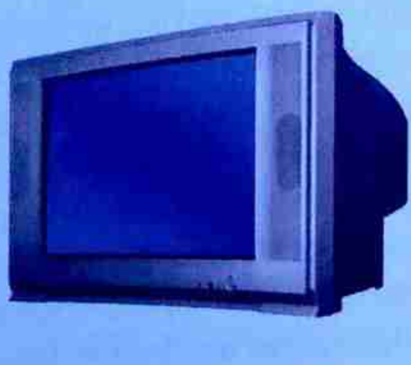


Figure 4.3: *CRT based TV set*

When only one type of phosphor is used as the screen, the image formed by the TV is black and white. Such a TV is known as a black and white TV. In a coloured television, there are three different phosphors, one for each primary colour; red, green and blue. The image is formed by varying the intensity of the electron beam that strikes the different phosphors. Some coloured televisions use a single electron gun whereas others use three.

Computer displays work in the same way as the television. However, CRT-based computer monitors and televisions are no longer in use due to the development of new technologies. These new technologies include Liquid Crystal Display (LCD) and Light Emitting Diode (LED). Both LCD and LED are cheap and less complex technologies compared to the CRT based display.

The cathode ray oscilloscope

Another important application of the CRT, is the Cathode Ray Oscilloscope (CRO). This is a versatile laboratory instrument which is very useful in displaying, measuring and analysing waveforms. The CRO is also useful in studying time varying quantities. The working principle of CRO is similar to that of CRT-based television set. The signal to be investigated is applied to the Y-plate of the CRO. The plate therefore, deflects an electron beam vertically. A sweep voltage is applied to the X-plates which causes an electron beam to be deflected horizontally at a constant rate. This makes a horizontal line which is the reference of measurement (time base). When the two beams are displayed on the screen, the waveform corresponding to the input signal can be observed. Figure 4.4 shows a CRO with two waveforms displayed on the screen.

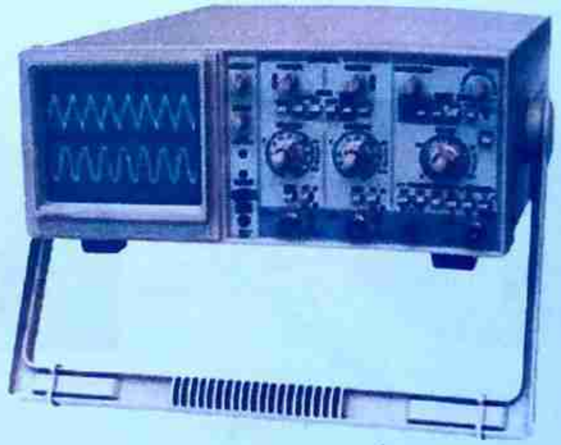


Figure 4.4: *Display of waveforms on a CRO screen*

X-rays

When an atom absorbs energy, some of its electrons in the inner shells may get enough energy to raise to the outer shells. An electron in an outer shell is at higher energy level than an electron in an inner shell. If an electron rises to an outer shell, it leaves a vacancy in its original inner shell. This makes the atom to re-arrange its electrons so that the vacancy in the inner shell is filled by an electron from one of the outer shells. In doing so, the electron loses part of its energy in order to have energy equal to the energy of the shell with a vacancy. The energy that is lost by the electron is emitted from the atom in the form of electromagnetic radiation, named X-rays. These rays are therefore, similar to visible light but have much higher frequency ranging from 3×10^{17} Hz to 3×10^{19} Hz. Note that, visible light frequency range is

4×10^{14} to 8×10^{14} Hz. X-rays have various uses and can be produced using a device known as the X-ray tube.

The X-ray tube

The X-ray tube consists of a vacuated glass, a wire filament, a target material, a copper anode and a beryllium window. The wire filament (normally tungsten wire) is connected to a low voltage power supply which heats the filament. The hot filament emits electrons through the thermionic emission. A high voltage supply is connected between the filament cathode and a copper anode. The potential difference (p.d) between the anode and the cathode accelerates the emitted electrons towards the copper anode. Upon arriving at the anode, the accelerated electrons strike the target which is embedded to the anode. The target is normally tungsten or molybdenum. When electrons strike the target, most of their energy is turned into heat energy. However, some energy causes the re-arrangement of the electrons in the atoms. This results into emission of X-rays which pass out of the tube through a beryllium window. The heat generated at the copper anode is carried away by a coolant such as oil or water which circulates around the anode. The whole X-ray tube is covered by lead so as to protect users of the device from exposure to X-rays. Figure 4.5 illustrates the structure of an X-ray tube.

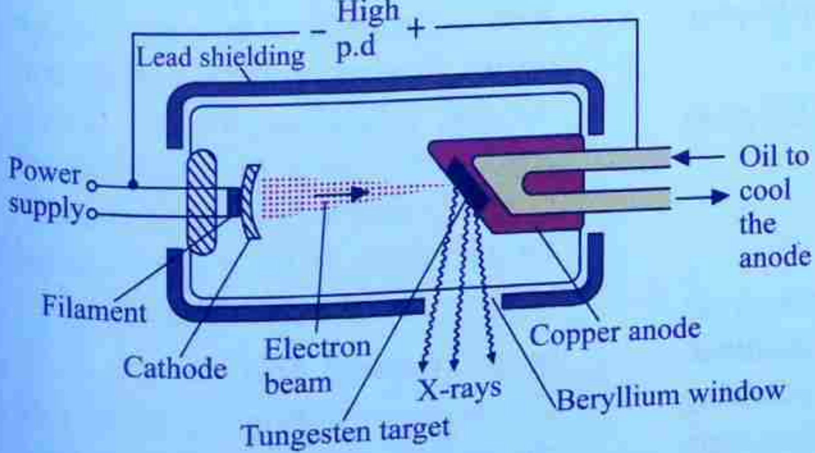


Figure 4.5: *Structure of an X-ray tube*

Production of X-rays

Electrons emitted by thermionic emission from the filament experience the potential difference and are accelerated towards the anode. When they hit the anode, the electrons are stopped and thereby transfer their energy to the electrons of the anode material. The anode material undergoes the re-arrangement of electrons resulting to the emission of X-rays. Only a small percentage of the incident electrons' energy is converted to X-rays. Large percentage of the electrons energy is lost as heat.

Task 4.3

Study the structure of the X-ray tube. Draw a well-labelled diagram of an X-ray tube and explain its working principle.

Hard and soft X-rays

The energy of the X-rays produced in the X-ray tube vary depending on the high voltage applied between the anode and the cathode. High p.d between the cathode and the anode leads to very energetic X-rays and low p.d leads to X-rays of low energy.

X-rays with high energy have short wavelengths while those of low energies have long wavelengths. X-rays of very short wavelengths have very high penetrating power and are in some cases referred to as hard X-rays. Conversely, X-rays with long wavelength have lower penetrating power and are sometimes referred to as soft X-rays.

Properties of X-rays

X-rays are electromagnetic radiation as are radio waves, gamma rays and others. Therefore, X-rays have similar properties to those of other electromagnetic waves. The following are some of the properties of X-rays:

(a) They travel in straight lines with the speed of light in vacuum, 3×10^{8} m/s.

(b) They can penetrate various objects and the degree of penetration depends upon their energy.

(c) They are not affected by electric or magnetic fields (they do not possess charge).

(d) They cause fluorescence upon striking certain types of crystals.

(e) They are invisible to the human eye, so their detection is possible only by means of detectors and photographic emulsions.

(f) They can ionise gases causing the gases to conduct electricity.

Applications of X-rays

X-rays have a wide range of applications in various fields including the medical, industrial and research fields. The following are some important fields where X-rays are very useful.

In the medical field

X-rays are used in hospitals to obtain diagnostic information and treatment of different diseases. In diagnostic procedure, X-rays are used to detect broken or fractured bones, and some diseases in soft tissues. Some notable examples are the chest X-ray images like the one in Figure 4.6. The chest X-ray image can be used to identify lung diseases such as pneumonia and lung cancer. They are also used in cancer treatment (radiotherapy) to destroy cancer cells.

Figure 4.6: *X-ray image of human chest*

In the industrial field

X-rays are used in non-destructive testing to inspect metal casting and welded joints for hidden faults. Also, industrial X-ray machines are primarily used to find toxic material in food products and to inspect luggage at airports and building entries.

In the research field

Researchers use X-rays to investigate the atomic structure of crystalline materials and obtain elemental information from different types of materials. Also, distant galaxies can be observed by detecting the X-rays emitted from the galaxies.

Project

Visit a radiology department of a nearby hospital with your teacher. Observe the X-ray production equipment and ask relevant questions regarding its functions and applications.

Write a report on the production and uses of X-rays after your visit. The report should also include the dangers posed by X-rays and prevention measures.

Chapter summary

1. Cathode rays are beams of accelerated electrons that form images upon striking a fluorescent screen.
2. The cathode ray tube is used in televisions, computer displays and cathode ray oscilloscopes.
3. X-rays are produced when fast-moving electrons strike a target and lose their energy.
4. Hard X-rays have shorter wavelengths and higher energy leading to high penetrating ability.
5. Soft X-rays have longer wavelengths and less energy leading to low penetrating ability.
6. X-rays are used in medical field, astronomy, security, crystallography, X-ray microscopic analysis, X-ray fluorescence, and industries.

Revision exercise 4

Choose the most correct answer in items 1 to 7.

1. What is the name of a process of ejecting electrons from hot metal surfaces?
   (a) Photoelectric effect.
   (b) Thermionic emission.
   (c) Static emission.
   (d) Thermal emission.
2. Why is vacuum needed in the cathode ray tube?
   (a) It increases the number of electrons.
   (b) It impedes or slows down the electrons.
   (c) It allows the electrons to flow unimpeded.
   (d) It allows the electrons to be easily scattered.
3. What happens when cathode rays strike a fluorescent screen?
   (a) Electrons are accelerated.
   (b) Nothing happens.
   (c) The screen emits light.
   (d) An electric current is produced.
4. How can the rate of thermionic emission from a metal surface be increased?
   (a) By immersing the metal in an electrolyte.
   (b) By increasing the surrounding pressure.
   (c) By decreasing the temperature of the metal to stabilize electron.
   (d) By increasing the surface area of the metal.

5. What is the purpose of connecting the anode of a CRT to a high voltage d.c power supply?
   (a) To deflect or change the direction of the electron.
   (b) To accelerate the electron.
   (c) To ionise gases.
   (d) To increase the rate of thermionic emission.
6. What are X-plates in a CRT?
   (a) Parallel plates positioned vertically to deflect the electron beam in a horizontal direction.
   (b) Parallel plates positioned horizontally to deflect the electron beam in a vertical direction.
   (c) Parallel plates positioned to focus the beam of electrons vertically.
   (d) Parallel plates set to create up and down motion of the electron beam.
7. What do you need to control, if you want to adjust the brightness of the signal in a CRO?
   (a) Focusing anode voltage.
   (b) Grid voltage.
   (c) Accelerating anode current.
   (d) Cathode size.
8. (a) What are the evidences that cathode rays are negatively charged particles?
   (b) With the aid of a labelled diagram, describe the functions of each part of the CRT.
   (c) Explain how a CRT differs from CRO.
   (d) Describe how cathode rays are produced in a cathode-ray tube.
   (e) Give *five* properties of cathode rays.
9. (a) Draw a well-labeled diagram of an X-ray tube.
   (b) Explain how X-rays are produced.
   (c) Give *five* properties of X-rays.
   (d) List any *five* applications of X-rays.
10. Distinguish between hard and soft X-rays. Explain the use of each.
11. In what ways do X-rays differ from light rays?
12. Figure 4.7 shows a section of a cathode ray oscilloscope.

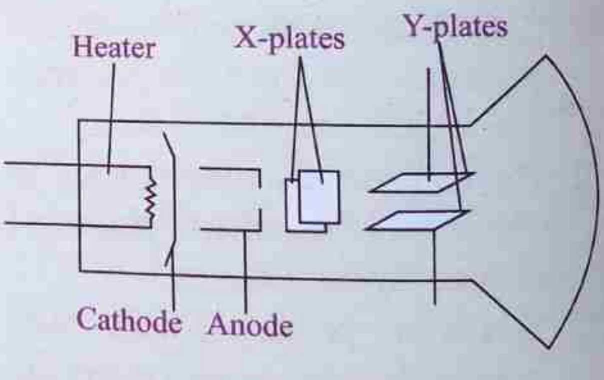


Figure 4.7

What changes should be made in order to produce the following on the screen?
   (a) A brighter trace
   (b) A vertical line
   (c) A wave pattern
   (d) A horizontal line.
13. Describe the behaviour of cathode rays and X-rays when subjected to both magnetic and electric fields.

Chapter Five

Electronics

Introduction

Electronics is fundamental to modern life. The discovery of electrons was a landmark in physics that led to great technological advances. This is evident in an exceptional broad range of technologies which have made a great impact in our daily life. These technologies include mobile phones, computers, radios, televisions, cameras, and satellite dishes to mention a few. In this chapter, you will learn about concept of electronics, band theory in solid materials, classification of solid materials in terms of conductivity, semiconductors, diodes, transistors and electronic amplifiers. The competencies developed will enable you to demonstrate basic skills in electronic components and use them to construct and repair electronic devices.

Concept of electronics

The term 'electronics' originates from the word electron and it refers to the study dealing with the theory and use of devices in which electrons travel through vacuum, gas, or semiconductor medium.

Electronics differs from current electricity since it deals with the motion of electrons in semiconductors under the influence of applied electric and/ or magnetic fields while current electricity deals with the motion of electrons in conductors (metals). Electronics involves the application of fundamental nature of electrons and the way the motion of these particles could be utilized. The aspects of electronics also involve the emission of electrons, storage of electrical charges and amplification of electrical signals.

***Electronics** refers to emission, behaviour and effects of electrons or flow of electrons in semiconductors, vacuum or gas and devices.*

Application of electronics

In the modern technological world, electronics is a valuable and ever growing sector. There are plenty of fields where electronics is applied as follows:

1. **Consumer electronics**

   Electronics knowledge is widely used in office and domestic gadgets such as computers, calculators, scanners, printers, fax machines and projectors. Audio and video systems like headphones, Digital Versatile Disc (DVD) players, Video Cassette Recorders (VCRs), microphones, televisions, washing machines, loudspeakers, and videogame consoles are also examples of consumer electronics. These are widely used in our daily activities.

2. **Electronic automation in manufacturing industry**

In industries, automated machines have improved productivity and efficiency, resulting in reducing time and cost. The machines are also safe to utilize in such a way that they make difficult and dangerous works to become manageable. Therefore, in many industries, delegating some tasks to the automation systems have become a preferred option. Automation systems are mostly the products of control electronics.

3. **In telecommunication**

The world is continuously changing, and technology is at the centre of that transformation. A large number of people now own smartphones that are capable of performing various functions. These phones comprise of electronic components. Moreover, computers are increasingly used today for various applications including online communication. Electronics is advancing at a rapid pace, allowing humans to create novel products and contribute to the transformation of society.

4. **In meteorology and oceanography**

Monitoring of the weather and climate is done using a large number of automatic weather stations and RADAR. Weather data are measured by different sensors and stored in a programmable data logger. Thus, electronics have become important aids in weather observational activities.

5. **In medical field**

Major transformation in the medical profession has been witnessed as a result of advanced medical electronic instruments for control, data recording and physiological analysis. These instruments are mainly used for diagnosing, treatment, planning system and radiation treatment. Some medical electronics instruments include digital thermometers, Computerized Tomography (CT) scanners, ultrasound machines, conventional X-rays machines and Magnetic Resonance Imaging (MRI) scanners. Figure 5.1 shows some medical electronics instruments.

Conventional X-ray machine

CT scanner

MRI scanner

Figure 5.1: *Some medical electronic instruments*

6. Electronic automation in financial systems

An electronic payment system is a system that allows one to make transactions or pay for goods and social economic services without using cheques or cash. The system is also known as an online payment system or an electronic payment system. Examples of the electronic payment system are Automated Teller Machine (ATM), e-banking, simbanking and online shopping. The electronic payment system has improved financial control for institutions involved in internet-based banking and shopping worldwide.

Electronic circuit

Electronic devices are made up of circuits consisting of primarily active components supplemented with passive ones. Such circuits are described as electronic circuits and are used to perform a wide variety of tasks including: conversion and distribution of electric power; controlling and processing of signals and data. Thus, electronic circuits are composed of various components connected to each other. These components are mainly classified as active or passive components.

Active components require a source of energy in the form of direct current for their operation. Transistors, operational amplifiers and Integrated circuits (ICs) are examples of active components. Active components can control and amplify a signal. Figure 5.2 (a) shows some active electronic components. On the other hand, *Passive components* neither require external power source to operate nor amplify a signal. They depend on the energy provided from external circuits. Passive components include: resistors, capacitors, inductors, transducers, transformers, thermistors and switches. Figure 5.2 (b) shows some examples of passive electronic components.

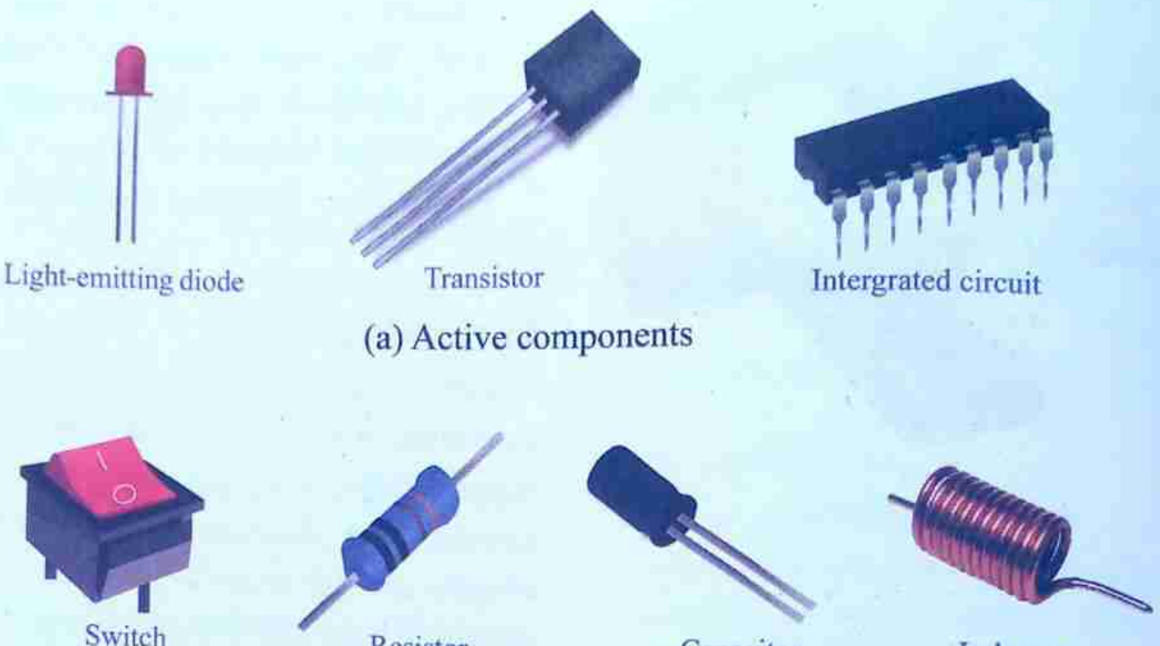


Figure 5.2: *Some electronic components*

Band theory in solid materials

The energy band structure (or simply band structure) of a solid is a series of "allowed" and "forbidden" energy bands that it contains. The series describes the ranges of energy band that electrons within a solid may occupy, and energy bands that electrons cannot occupy. The energy band that electrons can occupy are called *allowed energy bands* while those which cannot be occupied by electrons are called *forbidden energy bands*. The band structure determines the electronic properties of a solid. These bands include the valence band, conduction band and the forbidden energy gap.

The *valence band* is the highest energy band in an atom where electrons are normally present at absolute zero temperature. It is called valence band because it possesses valence electrons. This band can be completely or partially filled with electrons.

Above the valence band, exists a *Fermi level* which is the highest level that electrons can occupy at absolute zero temperature. The value of energy at the Fermi level is known as the Fermi energy. The Fermi level changes when a solid is warmed or the number of electrons in the solid changes.

The *conduction band* is the energy band above the valence band. It may be empty or partially filled with electrons. In conduction band, electrons are detached from the nucleus of the atom. Thus, electrons in this band can move freely and are therefore called conduction electrons. These electrons are responsible for the electrical conductivity of solids.

The *forbidden energy gap* is the band of energy separating the valence band and the conduction band. It is formed above the top of the valence band and below the bottom of the conduction band. In solids, electrons cannot stay in the forbidden band gap because, in this band, there is no allowed energy state. Figure 5.3 shows the arrangement of energy bands in a solid.

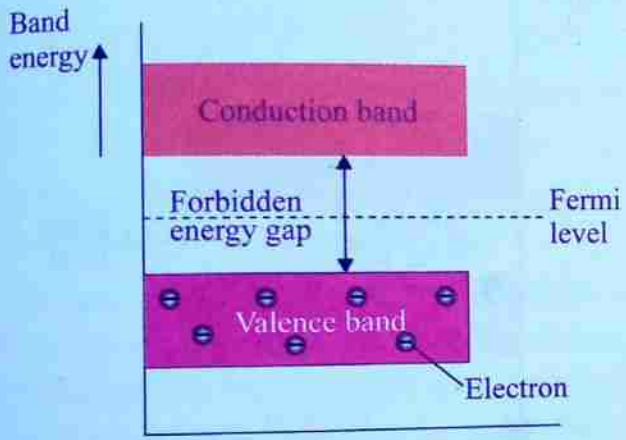


Figure 5.3: *Energy bands in a solid*

Classification of solid materials in terms of conductivity

Solid materials can be classified based on how large the forbidden gap is. When electrons move from the valence band to the conduction band, they have to cross through the forbidden energy gap. When the forbidden energy gap is large, electrons need large amount of energy (more than or equal to the energy gap) to cross the forbidden energy gap. If the forbidden energy gap is small, electrons need small amount of energy to cross the gap. If there is no forbidden gap between the valence and conduction bands, the two bands will come into contact or overlap such that electrons flow easily from the valence band to the conduction band. Therefore, the larger the forbidden energy gap, the more the energy required to transfer electrons from valence band to the conduction band. Based on the size of the forbidden energy band, solid materials can be classified as conductors, insulators and semiconductors.

Conductors are materials that have high electrical conductivity. In terms of the band theory, in conductors the valence band overlaps the conduction band with the Fermi level lying in the conduction band. That is, there is no forbidden band gap between the two, as shown in Figure 5.4. Due to this overlapping, many free electrons are also available in the conduction band and they are responsible for the conduction of electric current. Metals such as copper, gold, aluminium, silver and iron are examples of conductors.

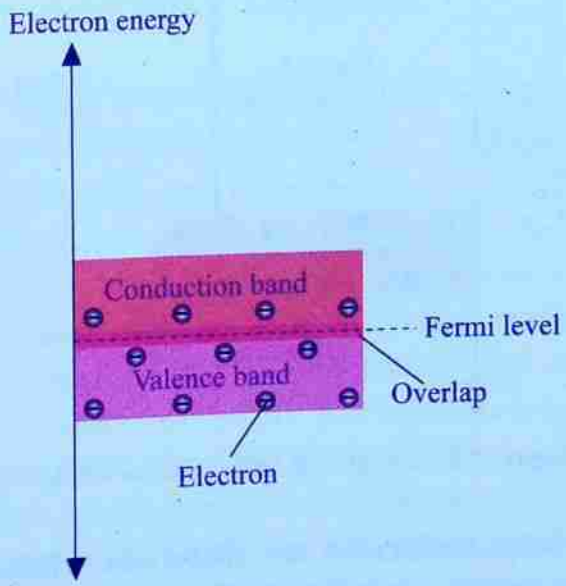


Figure 5.4: *Energy band structure for conductors*

Insulators are materials that do not conduct electric current. Most of the insulators are materials having very high electrical resistivity. In these materials, valence electrons are tightly bound to the atoms with no free conduction electrons. Considering the band theory, the valence band of insulators is full, whereas the conduction band is empty and the forbidden band is very large. The Fermi level in this case lies within the forbidden band. Therefore, a large amount of energy is needed to transfer the valence electrons from the valence band to the conduction band. That is why insulators under normal conditions do not conduct electric current. Examples of insulators are rubber, plastics, glass, mica, and quartz. Figure 5.5 shows the energy band structure for insulators.

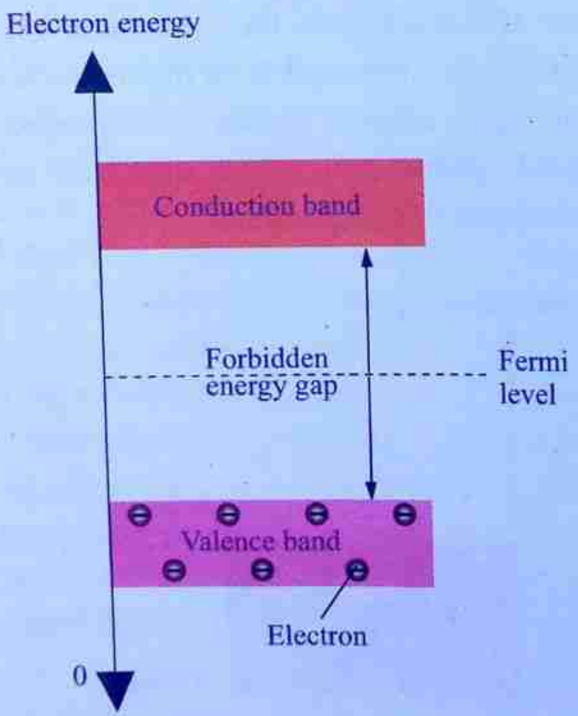


Figure 5.5: *Energy band structure for insulators*

Semiconductors are materials whose electrical conductivities lie between those of conductors and insulators. In their pure state, semiconductors are neither conductors nor insulators, and under certain conditions, they can conduct current electricity. Silicon is the most common material used to build semiconductor components. A semiconductor has a small forbidden energy gap between the valence and conduction bands as illustrated in Figure 5.6.

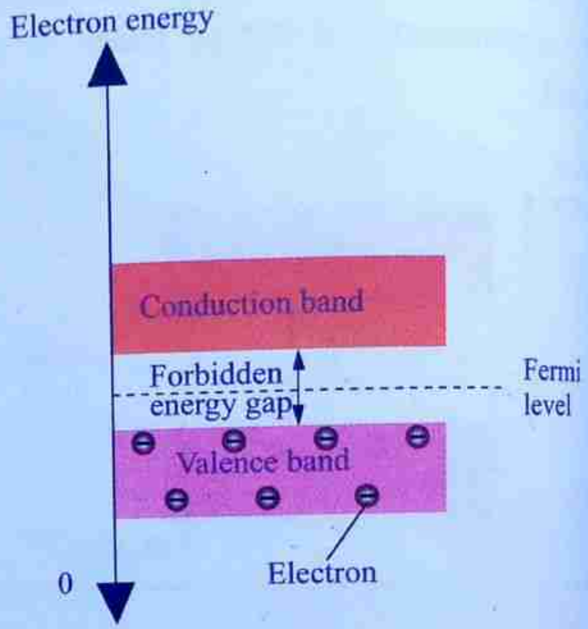


Figure 5.6: *Energy bands for semiconductors*

The Fermi level in semiconductors is also located within the forbidden energy gap. Therefore, a very small energy is needed to transfer the valence electrons from the valence band to the conduction band. The required energy is smaller than the required energy in insulators but larger than the required energy in conductors. Examples of semiconductors are:

1. Single element semiconductor such as silicon, boron, germanium, arsenic, and antimony.
2. Compound semiconductors such as gallium arsenide, indium phosphide, gallium nitride, silicon carbide, silicon germanium, zinc oxide and titanium oxide.

The energy band structures for conductors, semiconductors and insulators are compared in Figure 5.7.

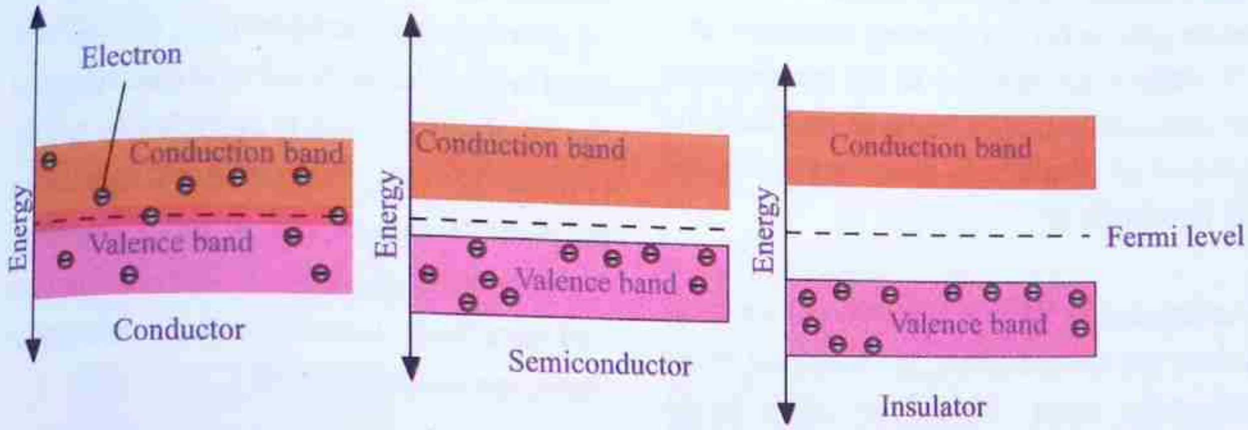


Figure 5.7: *Energy band structures for conductors, semiconductor and insulator*

Effect of temperature on electrical conductivity of solid materials

The amount and movement of free electrons in a material is affected by its temperature. A change in temperature in a material may affect the amount or movement of free electrons, or both amount and movement depending on whether the material is a conductor, a semiconductor or an insulator. Since the conductivity of conductors, semiconductors and insulators depends on the amount and movement of electrons it will consequently be affected by temperature as explained hereby.

Effect of temperature on the conductivity of conductors

In conductors, the conduction band overlaps with the valence band. As a result, a large number of free electrons are found in the conduction band. Since there are enough electrons in the conduction band, conductors can readily conduct electric current even at room temperature. With the increase in temperature, the electrical resistance of metals increases. An increase in resistance is due to increased vibrations of atoms with the increase in temperature, leading to increased collisions between the vibrating atoms and the moving electrons. So, despite having greater kinetic energy at higher temperature, the electrons face much more hindrance in their path from the vibrating atoms. This hindrance in the motion of electrons is the reason for the increased resistivity of the conductor, and hence decreased conductivity.

Effect of temperature on the conductivity of semiconductors and insulators

At absolute zero temperature, electrons in semiconductors are held tightly by their atoms, and no free electrons are present in the conduction band. Therefore, at this temperature, the semiconductor behaves like a perfect insulator. As the temperature increases, electrons in the valence band gain kinetic energy and break away from their shells. Since semiconductors have a narrow forbidden energy gap, electrons gain enough energy and jump to the conduction band, becoming conduction electrons responsible for the conduction of electric current. The transfer of the electrons from the valence band to the conduction band increases with increasing temperature. Therefore, an increase in the temperature of the semiconductor material increases the amount of conduction electrons and hence its conductivity.

For insulators the conductivity is increased when the temperature is increased from absolute zero. However very large temperature may be needed to enable some of the valence electrons to break away from valence to conduction bands and hence conduct electricity. Figure 5.8 shows the effect of temperature on conductivity of materials.

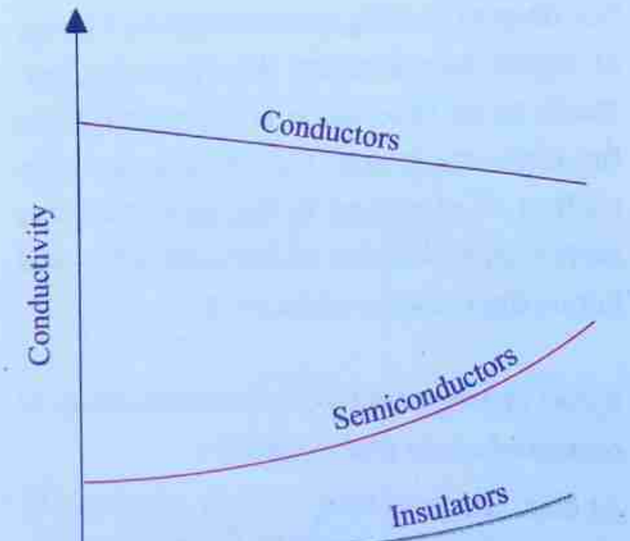


Figure 5.8: *Effect of temperature on the conductivity of material*

Types of semiconductors

Depending on the purity, semiconductors are of two types namely *intrinsic* and *extrinsic* semiconductors.

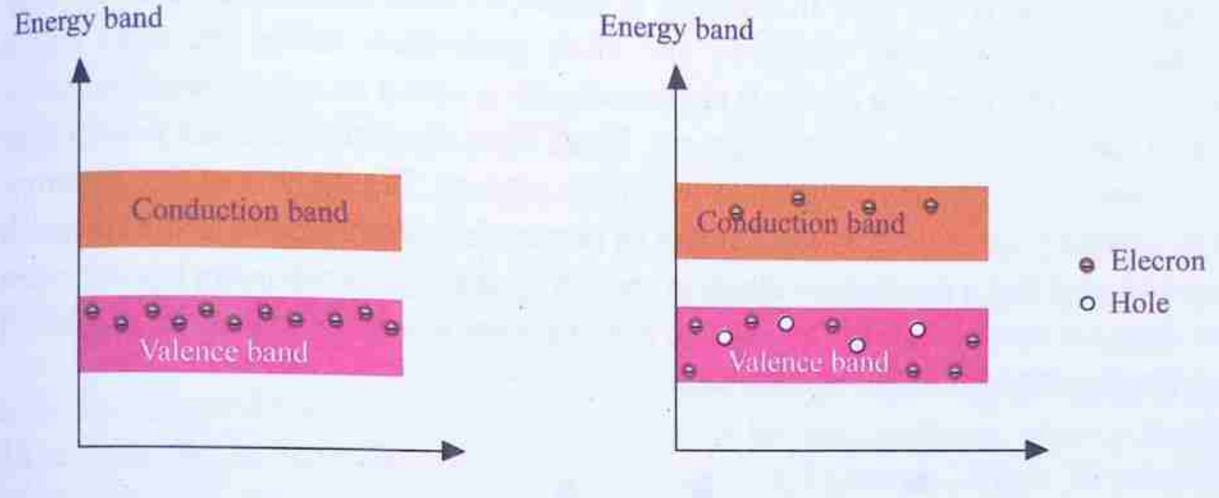
Intrinsic semiconductors

An intrinsic semiconductor is a semiconductor which is extremely pure. At absolute zero temperature, the valence band is completely filled with electrons and the conduction band is completely empty (Figure 5.9 (a)), making the material an insulator. Both pure silicon and germanium are examples of intrinsic semiconductors and have four electrons in their outermost shells, or valence shells.

As the temperature increases, the electrons gain more energy and break away from their shells becoming free electrons contributing to conduction. The space from which an electron gets removed becomes an electron vacancy or a "hole" as shown in Figure 5.9 (b). A hole carries a positive charge. A valence electron can move into a nearby hole, leaving a hole where it came from. In this manner, a hole appears to move from one site to another. Thus, the conduction in an intrinsic semiconductor is either due to the movement of free electrons in the conduction band or holes in the valency band.

In intrinsic semiconductors, the number of free electrons is equal to the number of holes as shown in Figure 5.9 (b). The hole may serve as a carrier of electricity whose effectiveness is comparable to that of a free electron. Electrons and holes in the conduction and valence bands, respectively, are referred to as charge carriers.

The conductivity of intrinsic semiconductors depends on the operational temperature. However, the conductivity of these semiconductors at room temperature is very low. This is because only few free electrons are present in the conduction band and few holes are in the valence band.



(a) At absolute zero temperature (b) Above absolute zero temperature

Figure 5.9: *Holes and free electrons in an intrinsic semiconductor*

Extrinsic semiconductors

Silicon and germanium in their pure form must be modified to increase their electrical conductivity by increasing the number of free electrons or holes in order to make them suitable for developing electronic components. This can be achieved by making use of suitable impurities. The deliberate addition of impurities to a pure semiconductor is called *doping*, and the impurities are referred to as *dopants*. The materials formed by adding dopants are referred to as doped semiconductors or extrinsic semiconductors. An extrinsic semiconductor is therefore a semiconductor that has been doped with impurities to modify the number and type of charge carriers in the semiconductor. By doping a semiconductor, one can increase its conductivity by a significant factor. For instance, in modern integrated circuits, heavily doped silicon is often used to replace metals.

Mechanism of doping semiconductors

Doping has to be done such that it does not distort the structure of the original pure semiconductor. Dopants have to occupy only a few of the original semiconductor atom sites in the solid. To attain this, the sizes of dopants and the semiconductor atoms should be nearly the same. There are two types of dopants: pentavalent and trivalent (Group III) elements in the periodic table. The former gives an *n*-type semiconductor, and the latter produces a *p*-type semiconductor.

n-type semiconductors

A pentavalent atom is added to increase the number of free conduction electrons in an intrinsic semiconductor. Pentavalent atoms are those with five valence electrons, including phosphorus (P), bismuth (Bi), arsenic (As), and antimony (Sb). Doping silicon with a pentavalent atom, such as antimony, forms covalent bonds whereby the antimony shares its four electrons with four adjacent silicon atoms. This leaves one extra electron as a free electron as shown in Figure 5.10.

The electron that is not used in bonding becomes a conduction electron. Since they give up an extra electron, antimony and other pentavalent atoms are often called *donor* atoms. The resulting extrinsic semiconductor is called an n-type semiconductor, where n stands for excess negative charge. Since there are extra electrons in an n-type semiconductor, electrons are the majority charge carriers. The number of free electrons can be carefully controlled by the number of pentavalent atoms added to the silicon. It is good to note that a conduction electron created by doping silicon using a pentavalent atom does not leave a hole in the valence band because it is an excess to the number of electrons required to fill the valence band.

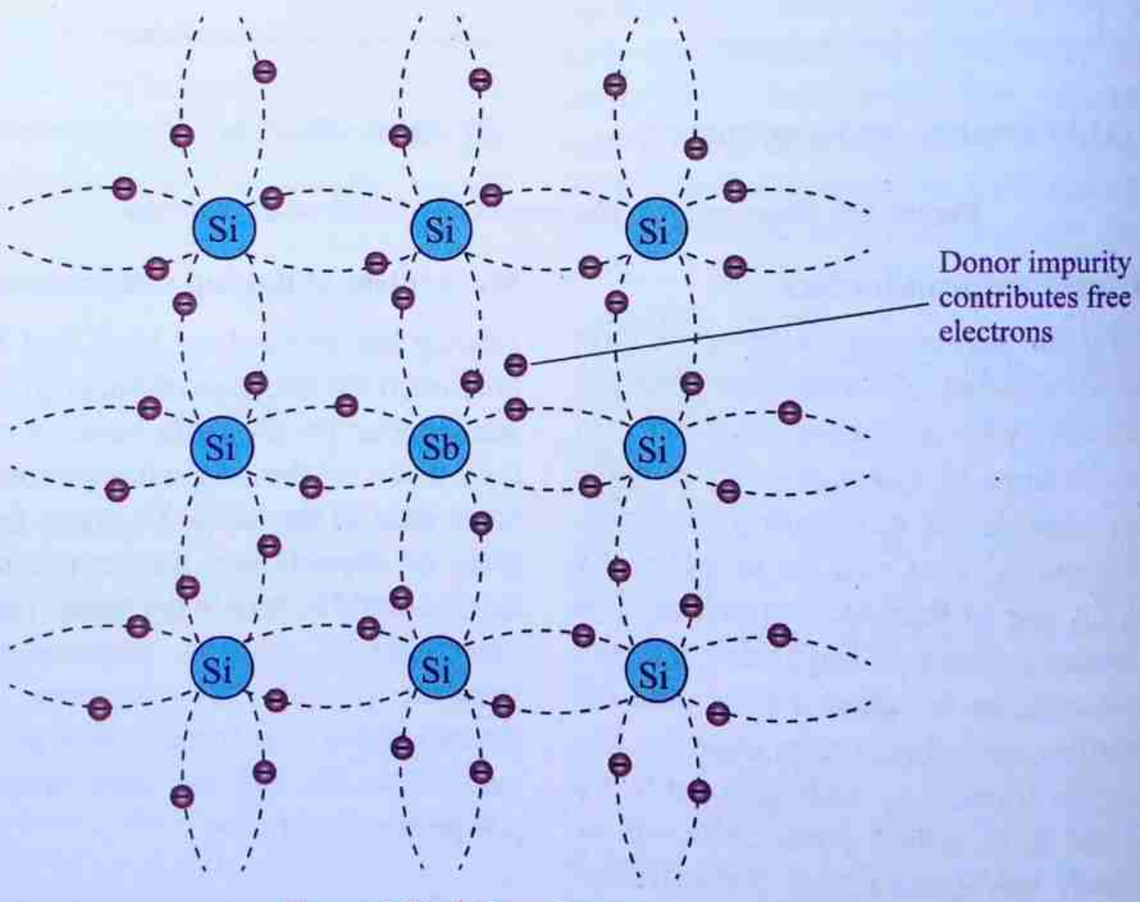


Figure 5.10: *Silicon doped with antinomy*

Although most charge carriers in n-type semiconductors are electrons, there are few holes present. The resulting holes are not produced by adding the pentavalent impurity atoms but are created when electron-hole pairs are thermally generated. Thus, the holes in an n-type material are called minority charge carriers. When an applied voltage sets the free electrons and holes in motion, electrons drift towards the positive terminal and holes to negative terminal. Figure 5.11 shows the direction of the flow of electrons and holes in an n-type semiconductor.

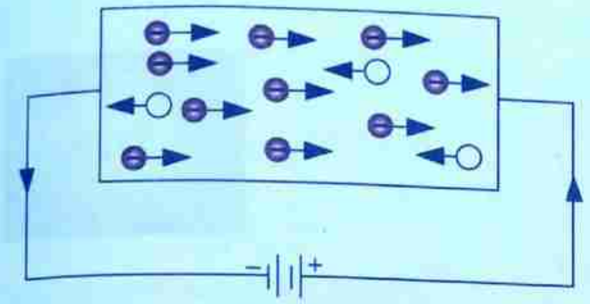


Figure 5.11: *Conduction in an n-type semiconductor*

p-type semiconductor

To increase the number of holes in an intrinsic semiconductor, trivalent atoms like boron (B), indium (In), and gallium (Ga) are added. Trivalent atoms are those with three valence electrons. When doping silicon with a trivalent atom such as boron, the atom forms covalent bonds with four adjacent silicon atoms. Since a boron atom has three valence electrons, a hole is formed, as all its three valence electrons are used to form three single covalent bonds with the host silicon atoms, which have four valence electrons as shown in Figure 5.12. Hence the fourth covalent bond is incomplete, having one electron less (vacant) since only silicon atom contributes an electron and boron has no electron to contribute to the bond. This missing electron (vacant) is called a *hole* and for each boron atom or any other trivalent element, a hole is created. Therefore, as it can take an extra electron, boron and other trivalent atoms are often called *acceptor* atoms. The resulting extrinsic semiconductor is called a *p*-type semiconductor, where *p* stands for the excess positive charge. Because there are extra holes in *p*-type semiconductors, holes are the majority charge carriers. The number of holes can be carefully controlled by the number of trivalent atoms added to silicon. It is good to note that a hole created by doping silicon using a trivalent atom does not accompany a free electron.

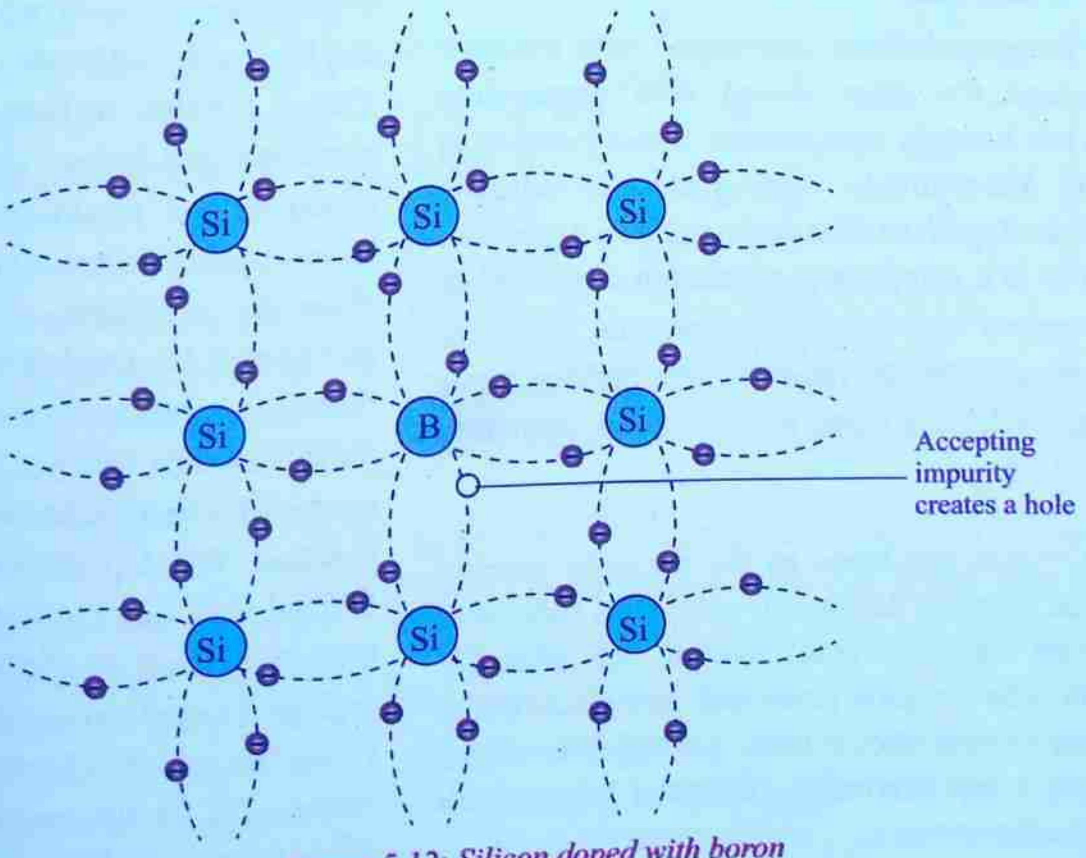


Figure 5.12: Silicon doped with boron

Although the majority charge carriers in p-type material are holes, a few conduction electrons are created when electron-hole pairs are thermally generated. Thus, electrons in p-type material are called minority carriers. When voltage is applied to a p-type semiconductor, holes diffuse towards the negative terminal and electrons towards the positive terminal, as illustrated in Figure 5.13.

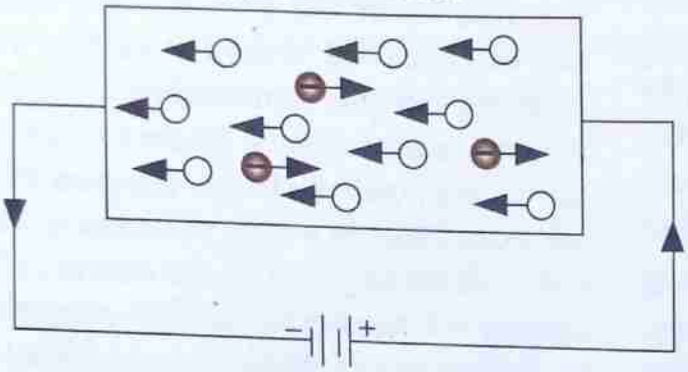


Figure 5.13: *Conduction in a p-type semiconductor*

The *p-n* junction

When semiconductors, one doped with trivalent atoms and the other doped with pentavalent atoms are brought into contact, a p-n junction is formed. Alternatively, when a piece of intrinsic silicon is doped so that one part is n-type, and the other is a p-type, a p-n junction is formed at the boundary between the two regions. The term *junction* refers to the region where the two types of semiconductors come into contact as shown in Figure 5.14.

The p region has holes as the majority carriers from the impurity atoms and only a few thermally generated free electrons which are minority carriers. The n region possesses free electrons as majority charge carrier from the impurity atoms and only a few thermally generated holes which are minority carriers.

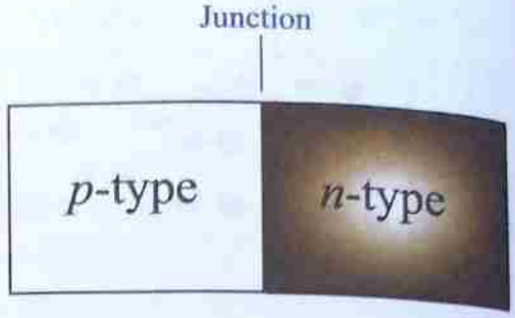


Figure 5.14: *p-n junction*

Mode of action of *p-n* junction

Once a p-n junction is formed, electrons from the *n-region* near the junction diffuse into the *p-region*, where they recombine with holes near the junction. This effectively creates a layer of positive ions in n-type region near the junction and a layer of negative ions in p-type region near the junction. A potential difference (p.d) is set up as the end of the n-side becomes positively charged and the p-side negatively charged. This p.d stops further flow of electrons and holes. These two positive and negative charge layers form the *depletion region* near the junction, as shown in Figure 5.15. Depletion region, therefore, refers to the region near the p-n junction depleted of charge carriers due to diffusion across the junction. The depletion region is formed very quickly and is very thin compared to the n and p regions. The p.d across this layer is called the *potential barrier* and is about 0.7 V for silicon and about, 0.3 V for germanium.

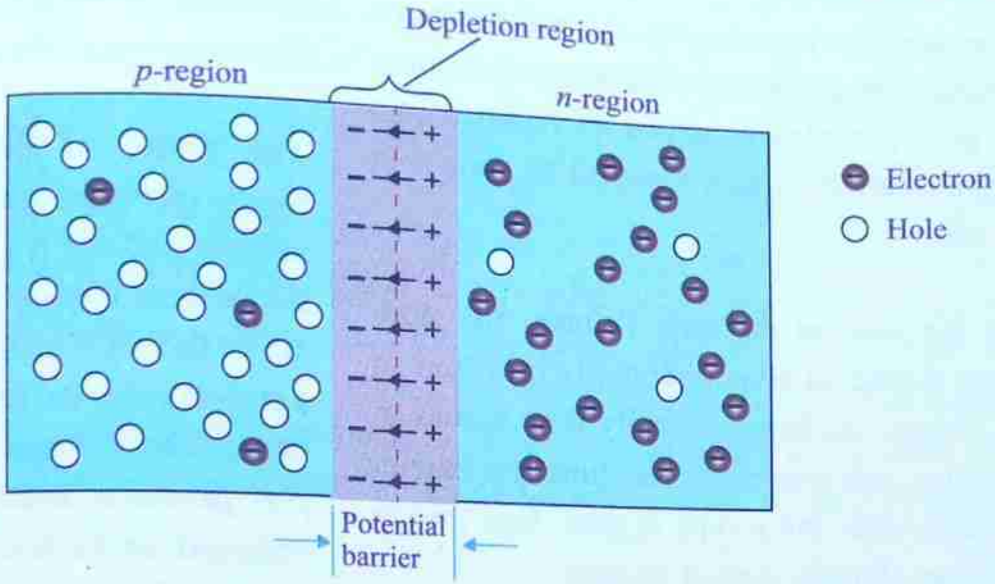


Figure 5.15: *Depletion region and potential barrier in a p-n junction*

Biasing of a *p-n* junction

When a potential difference is applied to a *p-n* junction, the junction is said to be under biasing. The applied p.d. is used to control the width of the depletion layer. Thus, biasing is a process of producing a set of currents or voltages at a *p-n* junction to form suitable operating conditions within an electronic component. There are two types of biasing a *p-n* junction or semiconductor junction. These are forward bias and reverse bias.

Forward bias of a *p-n* junction

Forward bias occurs when the *p*-type side of a *p-n* junction is connected to the positive terminal and the *n*-type side is connected to the negative terminal of a voltage source. That is, forward bias is the condition that allows electric current to flow through the *p-n* junction. Figure 5.16 shows a *p-n* junction connected to a voltage source in a forward bias mode.

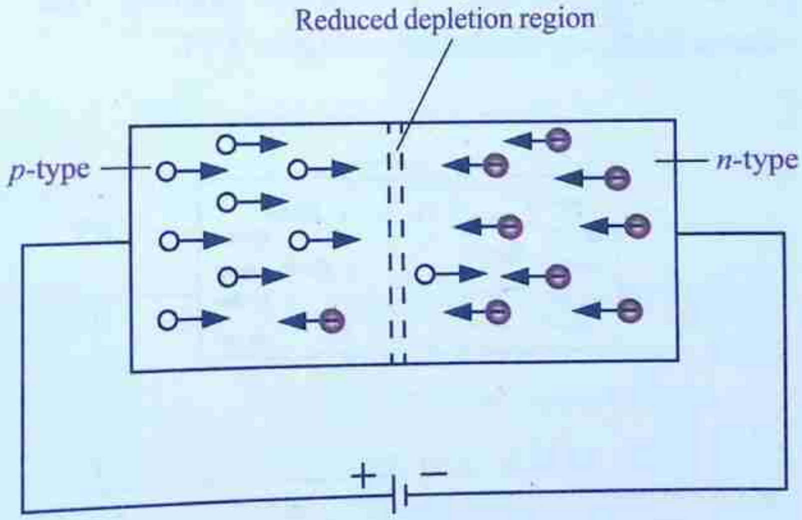


Figure 5.16: *Forward biased p-n junction*

In forward bias, the positive terminal connected to the *p*-type region repels the holes and pushes them towards the *p-n* junction, while the negative terminal connected to the *n*-type region repels the electrons and pushes them towards the *p-n* junction. The negative terminal of the voltage source also provides a continuous flow of electrons through the external circuit in the *n*-region. As the electrons and the holes are pushed towards the *p-n* junction, the size of the depletion region is reduced and the potential barrier is also lowered as shown in Figure 5.16

With the increase in the bias voltage, the non-conducting depletion layer eventually becomes so thin that charge carriers can easily flow across it. The electrons can overcome the junction barrier potential and enter the *p*-type region. This makes the flow of an electric current possible.

Reverse bias of a *p-n* junction

Connecting the *p*-type region to the negative terminal of a voltage source and the *n*-type region to the positive terminal produces a reverse bias effect. Reverse bias is the condition that essentially prevents current flow through the junction. The connections are illustrated in Figure 5.17.

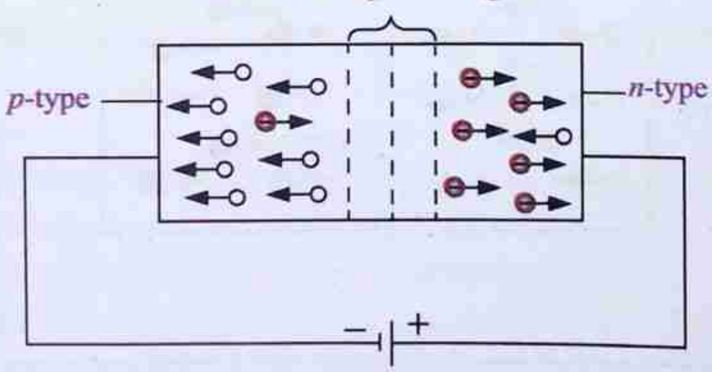


Figure 5.17: *Reverse biased p-n junction*

Because the *p*-type region is now connected to the negative terminal of the cell, the holes in the *p*-type region are pulled away from the *p-n* junction. Similarly, because the *n*-type region is connected to the positive terminal, the electrons will be pulled away from the junction. This effectively widens the depletion region and the junction offers a higher potential barrier. For this reason, there will be minimal or no electric current flowing through the *p-n* junction.

The width of the depletion region increases with increase in voltage and the p.d increases as the reverse voltage is increased. When the p.d is increased beyond a critical level, the junction breaks down. The voltage at which the *p-n* junction breaks down is referred to as the *breakdown voltage*. Junctions are intended to operate below their breakdown voltages. When the *p-n* junction is connected such that, it forms a two-terminal component, it is referred to as a *p-n* junction diode.

Current-Voltage characteristic curve for a *p-n* junction diode

The current-voltage characteristic curve, named as I-V characteristic curve, reflects the operation of the *p-n* junction. The *p-n* junction has a non-linear I-V characteristic curve unlike a resistor which has a linear I-V characteristic curve at room temperature. When a *p-n* junction is forward biased, a forward current will flow through it.

When the forward voltage reaches the internal barrier voltage of the *p-n* junction diode avalanche occurs, and the forward current increases significantly for a very small increase in voltage, leading to a non-linear curve indicated as forward curve's "knee" point in the first quadrant of the I-V characteristics as shown in Figure 5.18 (a).

When the *p-n* junction is reverse biased, it blocks current, serve for a very little leakage current. It operates in the third quadrant of its I-V characteristic curves as shown in Figure 5.18 (b). The diode continues to block current flow until the reverse voltage across it exceeds its breakdown voltage point, leading to a sharp spike in reverse current downward curve. Zener diodes make effective use of such a reverse breakdown voltage point. It is good to note that, the forward bias current is measured in milliampere (mA) while the reverse current in microampere (μA). Combining the I-V characteristic curves for the forward and reverse biased mode in one plot gives the graph shown in Figure 5.18 (c).

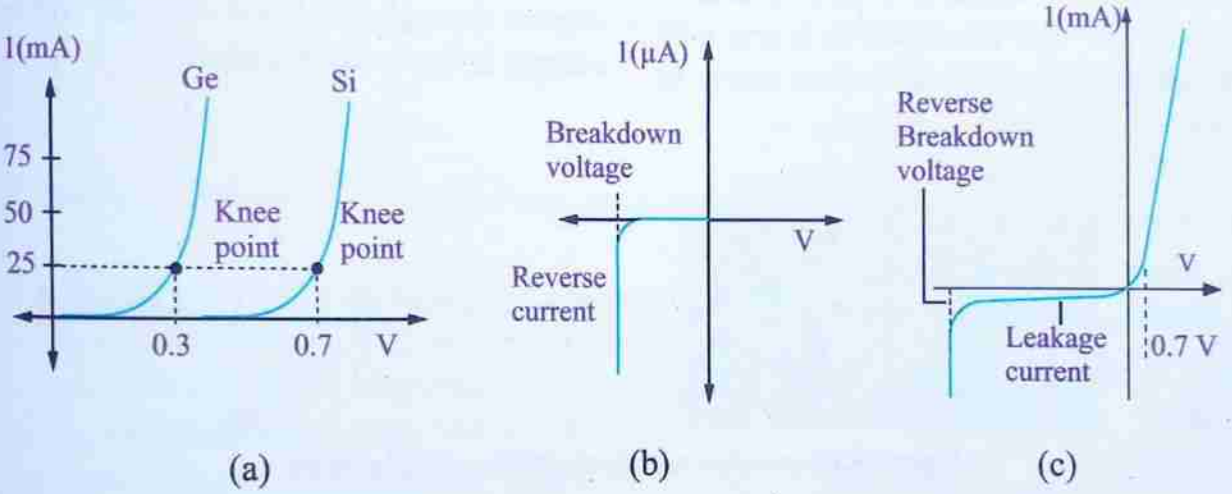


Figure 5.18: *I-V characteristic curve for biasing silicon p-n junction diode*

Exercise 5.1

1. What is the primary difference in the band structure of semiconductors as compared to insulators?
2. Why are there equal numbers of electrons and holes in an intrinsic semiconductor?
3. If a small number of impurities alters the electron or hole concentrations in an intrinsic semiconductor, how would this affect the electrical conductivity at low temperature?
4. How could electrical conductivity be used to determine whether a material is a conductor or semiconductor?
5. Classify the following components as either active or passive. Give reason(s) for the classification.

   Resistor, capacitor, inductor, integrated circuit, transistor, diode and zener diode.
6. Distinguish between electrical circuit and electronic circuit.
7. Give at least two differences between the following:
   (a) Intrinsic and extrinsic semiconductors.
   (b) *p*-type and *n*-type semiconductors.

Diodes

A diode is an electric device that allows the flow of current in only one direction. It is a single *p-n* junction device with electrical terminals connected in the *p*-type and *n*-type regions. The terminal of *p*-type region is called the anode and that of the *n*-type region is called the cathode. Such a diode is known as a *p-n* junction diode. Figure 5.19 (a) and (b) show a *p-n* junction diode and its symbol, respectively.

Therefore, the diode is a semiconductor component that functions as a one-way switch. It allows current to flow easily in one direction, but strictly limits current from flow in the opposite direction.

The property of a diode to allow current flow in one direction (forward bias) and restrict current flow in the opposite direction (reverse bias) is possible since the diode has a *p-n* junction and the depletion region between the two *p* and *n* semiconductor materials. When the *p-n* junction is forward biased, it allows current to flow from the anode to the cathode. The magnitude of the current through the diode depends on the current in the external circuit.

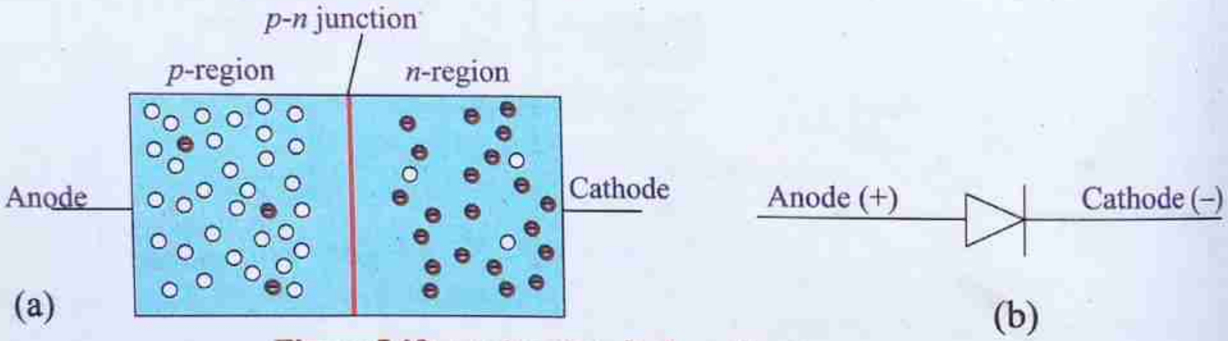


Figure 5.19: *p-n junction diode and its circuit symbol*

When the diode is reverse biased, extremely small current measured in microamperes flows through it. This current is known as reverse saturation current or reverse leakage current. It is the current that flows due to the minority charge carriers. This current is not high enough to make the diode conduct.

Types of diodes and their applications

There are different types of diodes used in electronic circuits. The following are the most common ones.

Semiconductor diode

One type of diode is the *p-n* junction diode as described in the previous section. This is also referred to as a semiconductor diode. Most semiconductor diodes are made up of silicon or germanium. Semiconductor diodes are used for very fast switching and microwave applications. Figure 5.20 shows a picture of a semiconductor diode and its symbol.

Figure 5.20: *Germanium diode*

Light Emitting Diode

A Light-Emitting Diode (LED) is a semiconductor diode that emits light when an electric current is applied in the forward direction of the diode. LEDs are made from a variety of semiconductor materials depending on the wavelength of the light required. The most commonly used materials for visible LEDs are gallium phosphide and gallium arsenic phosphide. Figure 5.21 shows LEDs and their curcuit symbol.

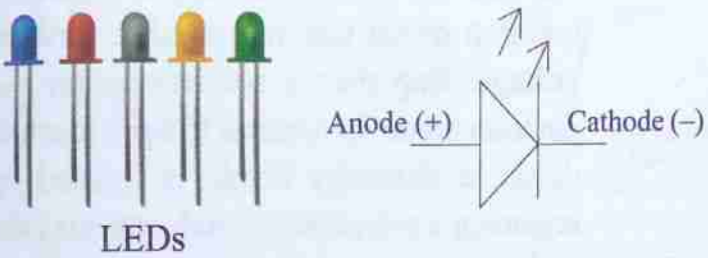


Figure 5.21: *Light Emitting Diodes and their circuit symbol*

LEDs have a wide range of applications; these include mobile phones, large advertising billboards, simple indicator lamps, large display screens, optical fibre communication links and as standard light sources in electrical equipment. Unlike traditional light sources, which convert only some of the electrical energy into light, the rest being wasted as heat. LEDs convert most of the electrical energy into light, resulting in efficient light production and hence minimal electric power consumption.

Zener diode

A Zener diode, shown in Figure 5.22 with its circuit symbol, is a semiconductor diode designed to be operated in the reverse breakdown voltage, called the Zener voltage. Every Zener diode is manufactured for specific reverse breakdown voltage.

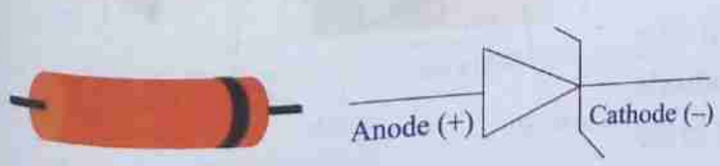


Figure 5.22: *Zener diode and its circuit symbol*

Zener diodes are among the main building blocks of electronic circuits and are widely used as voltage regulators in different electrical circuits.

Laser Diode

A laser diode is similar to LED because it converts electrical energy into light energy. Laser diodes, however, produce coherent light, as opposed to LEDs. This coherent light is generated using a process termed as "Light Amplification by Stimulated Emission of Radiation", abbreviated as LASER. Laser diodes are used in optical communication, laser pointer, CD drives, barcode readers and laser printer. Figure 5.23 shows a laser diode and the respective circuit symbol.

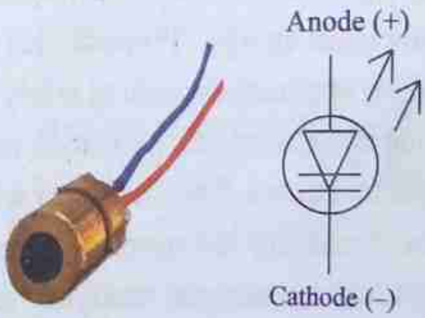


Figure 5.23: *Laser diode and its circuit symbol*

Avalanche Diode

An Avalanche diode is a *p-n* junction diode that operates in the avalanche breakdown zone. Avalanche breakdown occurs when enough reverse voltage is given to the *p-n* junction. As a result, the minority carrier ionizes and initiates a strong current flow in the reverse direction. Avalanche diode works electrically similar to the Zener diode. However, a Zener diode has a higher doping concentration than an Avalanche diode. Figure 5.24 shows Avalanche diodes and the circuit symbol.

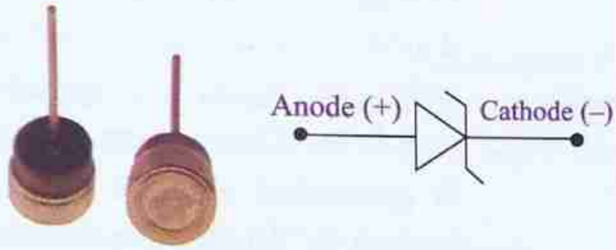


Figure 5.24: *Avalanche diodes and their circuit symbol*

Photodiode

A photodiode is a semiconductor device with a *p-n* junction that converts photons of light into electrical current. It is a type of light sensor that transforms light into electrical energy. Photodiodes are used in many applications such as safety electronics like fire and smoke detectors, cameras and photosensors. They are also widely used in a variety of medical applications, including sample analysis equipment, computed tomography detectors, and blood gas monitors. Figure 5.25 shows a photodiode and its circuit symbol. The photodiode symbol is similar to that of a LED, except the arrows in the photodiode point inwards rather than outwards as in the LED. It should be noted that a photodiode operates under reverse bias mode.

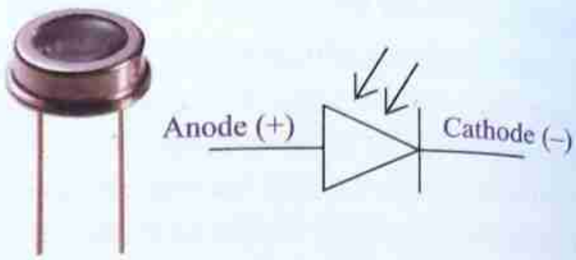


Figure 5.25: *Photodiode and its circuit symbol*

Schottky diode

The Schottky diode is a metal-semiconductor junction diode that has smaller forward voltage drop than a semiconductor *p-n* junction diode. In contrast to a *p-n* junction diode, a Schottky diode is formed by replacing a *p*-type semiconductor material with a metal like gold or platinum. When a metal is joined with *n*-type semiconductor, a junction is formed between the metal and *n*-type semiconductor, forming a metal-semiconductor junction diode called Schottky diode. The *n*-type semiconductor material acts as cathode terminal and the metal acts as the anode as illustrated in Figure 5.26. Schottky diodes provide small forward voltage drop and a very fast switching action. It is good to note that a Schottky diode conducts current in the forward direction when suffient forward voltage is applied. Schottky diodes are used as rectifiers and switching regulators.

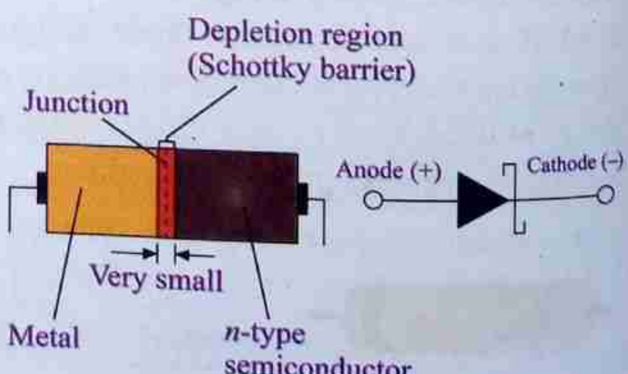


Figure 5.26: *Schottky diode and its circuit symbol*

The I-V characteristic curve of a Schottky diode is almost similar to that of a *p-n* junction diode except the knee voltage. A Schottky diode starts to conduct at much lower voltage compared to a *p-n* junction diode as illustrated in Figure 5.27.

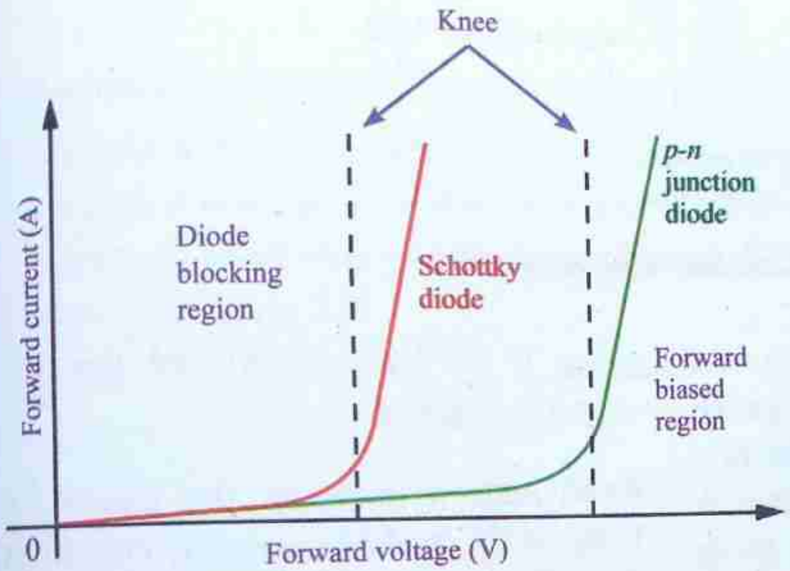


Figure 5.27: *I-V characteristics curve of Schottky diode*

Tunnel diode

A tunnel diode is a heavily doped *p-n* junction diode in which the electric current that flows through it decreases with an increase in voltage. In this diode, electric current is caused by tunneling. The tunnel diode is used as a very fast switching device in computers and in high frequency oscillators and amplifiers. A tunnel diode and its circuit symbol is shown in Figure 5.28.

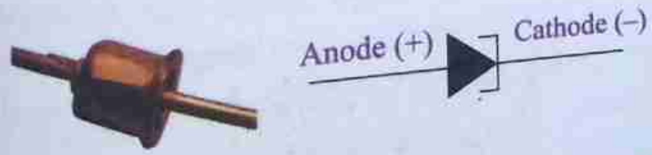


Figure 5.28: *Tunnel diode and its circuit symbol*

Rectification

Many electrical appliances and most electronic components require direct current (d.c) to operate. This means a d.c must be supplied to these devices. Such current may be provided by batteries. However, it is more convenient and economical to use an alternating current (a.c) than the d.c. power sources. Since the main supply is a.c, it must first be changed into d.c. The process of obtaining direct current from the alternating current is called *rectification*. The a.c rectification is one of the main applications of semiconductor diodes. Rectification can be done in two ways: *half-wave* and *full-wave* rectifications.

Half-wave rectification

In half-wave rectification, only one-half cycle of a.c input is converted to d.c. Figure 5.29 illustrates the half-wave rectification circuit. The input and output currents are both shown in the figure.

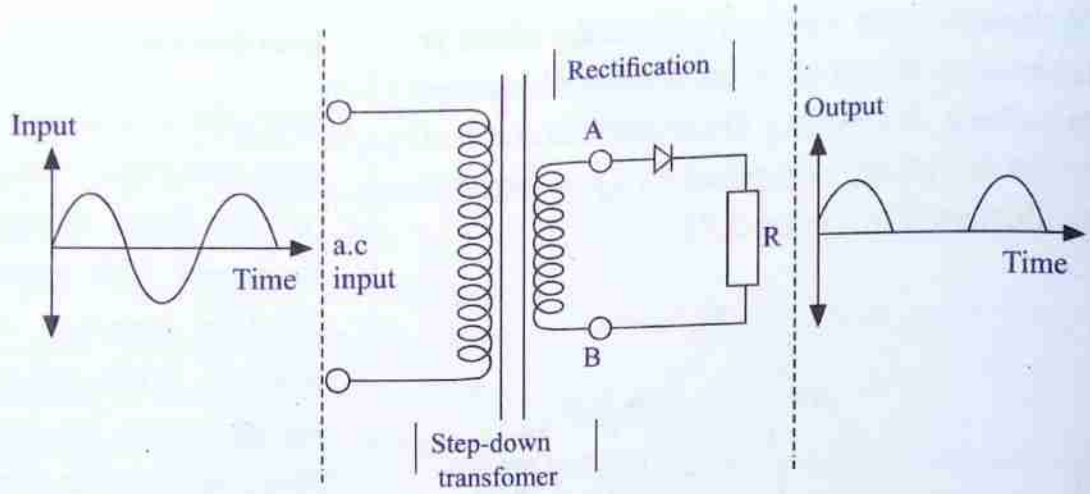


Figure 5.29: *Half-wave rectification*

When the current at A is increasing in a positive direction, it decreases in negative direction at B. Therefore, during the first half-cycle of the sinusoidal wave form, A is positive and B is negative. The diode is thus forward biased and current flows through the load resistor R.

During the second half cycle, A and the anode of a diode are negative, whereas B and the cathode of a diode are positive. The diode is thus reverse biased and the flow of current will be blocked. This is indicated by a horizontal line in the output diagram. The diode, therefore, conducts on every positive half-cycle and hence, half-wave rectification is achieved. The output is d.c and is always positive in value. If the diode is reversed, then the output is negative.

Rectification done by the circuit in Figure 5.29 results to unsteady current as the voltage keeps increasing from zero to maximum and then decreases to zero again. Therefore, the output current is not steady and it has to be smoothed, for it to be useful.

To smooth the output current, the rectification circuit can be modified by putting a large capacitor, C, in parallel with the load as shown in Figure 5.30. The capacitor is charged during the positive half-cycle of the a.c and discharges through the load in the negative half-cycle. This action is called *smoothing*.

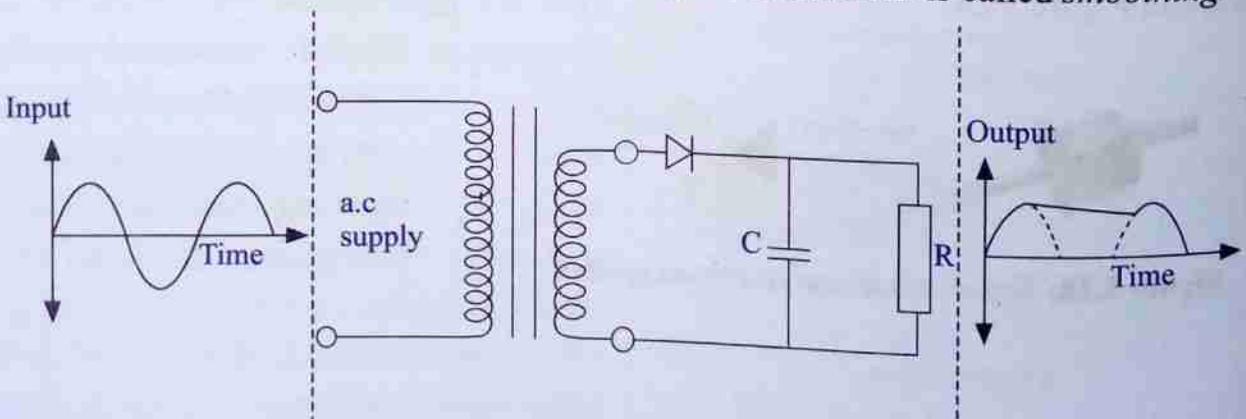


Figure 5.30: *Circuit for the half-wave rectifier with a smoothing capacitor*

The half-wave rectification with a smoothing capacitor is useful for low power applications. Yet, the method is not suitable for applications that need "steady and smooth" high d.c supply voltage. This is because in half-wave rectification, power is lost for every negative half of the cycle. This situation can be corrected by using every half cycle of the input voltage. This can be done by a circuit that does a full-wave rectification.

Full-wave rectification

In full-wave rectification, both halves of the a.c cycle are transmitted in the same direction. One way of achieving this is to have a transformer whose output has a centre-tap, that is, its output can be taken at two points one being half the other. The circuit and its output are shown in Figure 5.31.

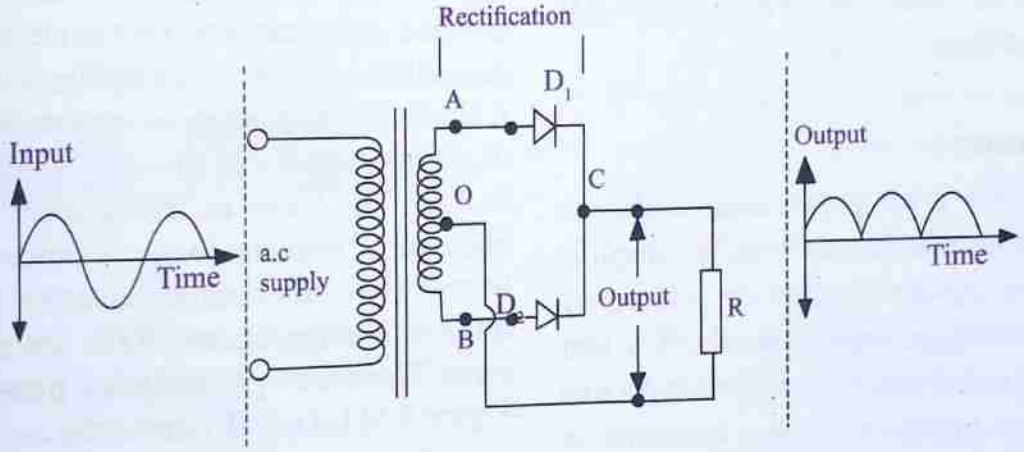


Figure 5.31: *Centre-tap full-wave rectification*

During the positive half cycle, point A is positive with respect to O and point B is negative with respect to O. Diode D_1 is forward biased but diode D_2 is reverse biased. The current passes through D_1, C, R and back to O. In the negative half-cycle, point B is positive with respect to O and point A is negative with respect to O. Diode D_2 is forward biased but diode D_1 is reverse-biased. The current passes through D_2, C, R and back to O. The direction of the current through R is the same as in the first half-cycle. The output is a full wave rectified voltage since both halves of each cycle of the a.c voltage appears in the d.c output.

Another method of achieving full-wave rectification is by using a bridge rectifier whose circuit is shown in Figure 5.32.

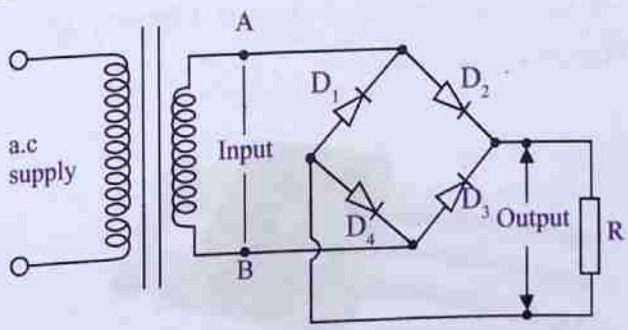


Figure 5.32: *Bridge rectifier for full wave rectification*

In the first half-cycle, point A is positive so that diode D_2 and D_4 are forward biased and D_1 and D_3 are reverse-biased. Diodes D_2 and D_4 conduct and the current flows from A via D_2, R, D_4 and back to the source at B. In the second half-cycle, B is positive and so diode D_3 and D_1 are forward-biased and D_4 and D_2 are reverse biased. Diodes D_3 and D_1 conduct and current flows from B via D_3, R, D_1, and back to the source at A. The current through R is in the same direction in both half-cycles. The output is the same as the one obtained using the centre tapped full wave rectifier.

Transistors

A transistor is an active semiconductor component, commonly used to amplify electronic signals (current or voltage) or switch electronic signals on or off. When two *p*-type and one *n*-type semiconductors are joined together, a *p-n-p* transistor is created. Likewise, when two *n*-type and one *p*-type semiconductors are joined together, an *n-p-n* transistor is formed. Transistors form the fundamental building blocks of modern electronic devices. It is used, for example, in amplifiers to make a sound signal stronger. Some transistors are packaged individually but most of them are found in integrated circuits. Figure 5.33 shows a transistor.

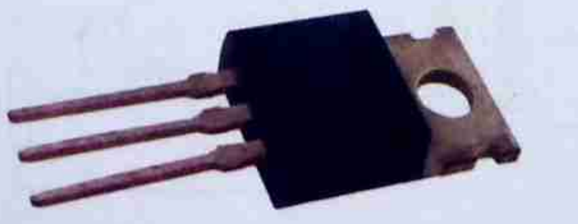


Figure 5.33: *Transistor*

Types of transistors

There are two broad categories of transistors. These are bipolar junction transistors (BJTs) and field-effect transistors (FETs). Bipolar transistors require a biasing input current at their control leads, and require both positive and negative charge carriers to operate. On the other hand, FETs require only voltage and one charge carrier to operate. This section introduces the BJTs only.

Bipolar junction transistors

Bipolar junction transistors are three-terminal components that act as electrically controlled switches or as amplifier controls. A bipolar junction transistor consists of a pair of *p-n* junction diodes that are joined back-to-back. This forms an arrangement where one type of semiconductor is sandwiched between two other semiconductors. Based on their configurations, BJTs are of two types. These are *n-p-n* and *p-n-p* transistors. Figure 5.34 (a) and (b) shows the *n-p-n* and *p-n-p* transistors and their circuit symbols.

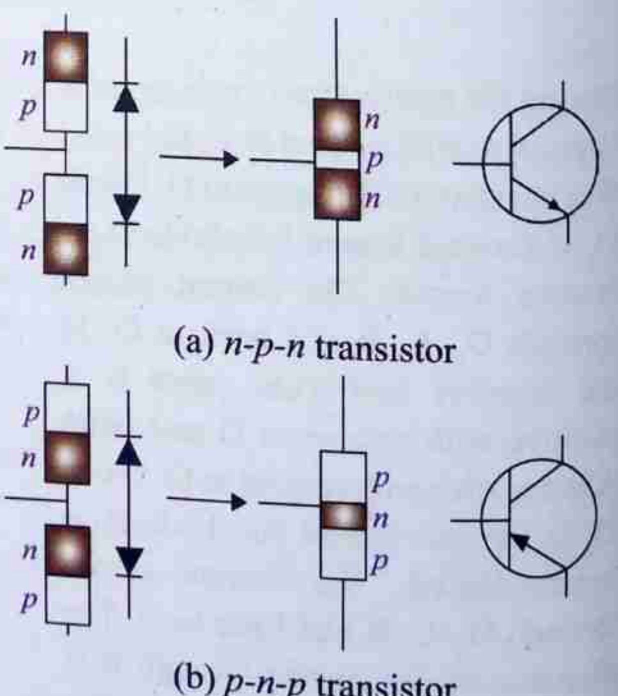


Figure 5.34: *Types of bipolar junction transistor and their circuit symbols*

The three terminals of the transistor are conventionally called the collector (C), base (B) and emitter (E) as shown in Figure 5.35. An easy way of identifying a transistor from its circuit symbol is using the direction of the arrow. The arrow indicates the direction of flow of conventional current. For a *p-n-p* transistor, the arrow points into the transistor. So, you can read *p-n* (for *p-n-p*) as "pointing in". For the *n-p-n* transistor the arrow points out (does not point in). So, you can read *n-p-n* as "not pointing in". Most transistors in use today are *n-p-n* because the majority charge carriers are electrons which are free and move faster than holes in *p-n-p*.

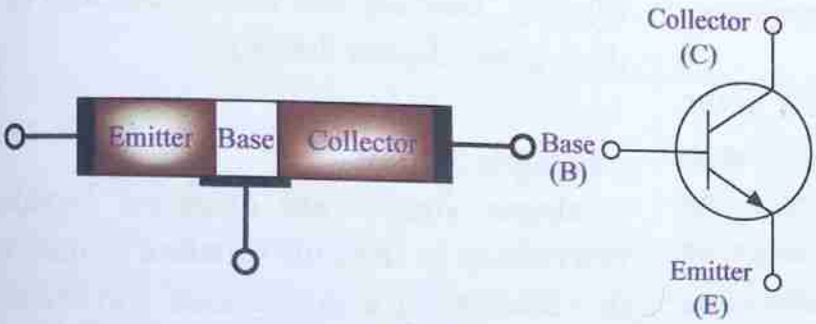


Figure 5.35: *n-p-n transistor terminals*

Mode of action of *n-p-n* transistors

To understand the mode of action of the transistor, consider the *n-p-n* transistor shown in Figure 5.36. The transistor is forward biased for emitter-base junction and reverse biased for collector-base junction.

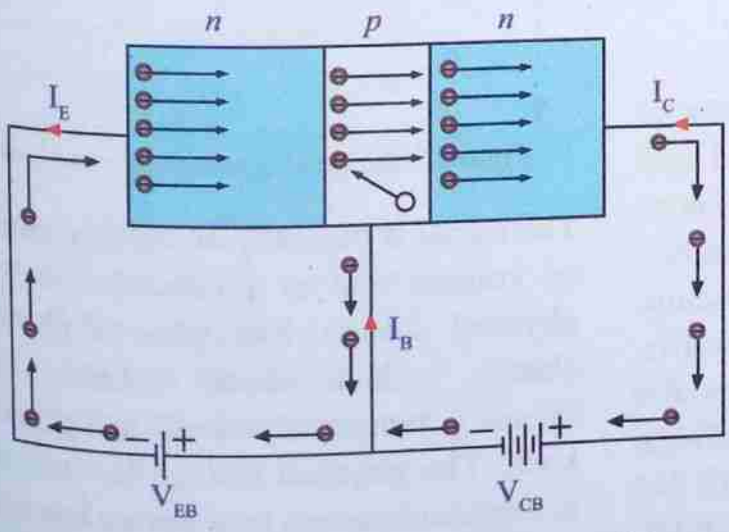


Figure 5.36: *Mode of action of n-p-n transistors*

The function of this transistor is based on the working principle of a *p-n* junction diode.

The electrons in the *n*-type emitter flow towards the base since it is forward biased. This results into flow of emitter current, I_E. It is conventional to consider the direction of the current to be in the opposite direction to the flow of electrons. These electrons tend to combine with holes when they pass through the *p*-type base. Because the base is narrow and lightly doped, only a few electrons couple with holes to form base current, I_B. More than 95% of the electrons cross into the collector region to form collector current, I_C. Thus, the total of collector and base currents form the emitter current, I_E. That is;

$$I_E = I_B + I_C$$

Mode of action of *p-n-p* transistor

The mode of action of a *p-n-p* transistor can similarly be shown using the circuit in Figure 5.37. The transistor is forward biased at the emitter-base junction and reverse biased at the collector-base junction. This transistor also, works based on the principle of a *p-n* junction diode.

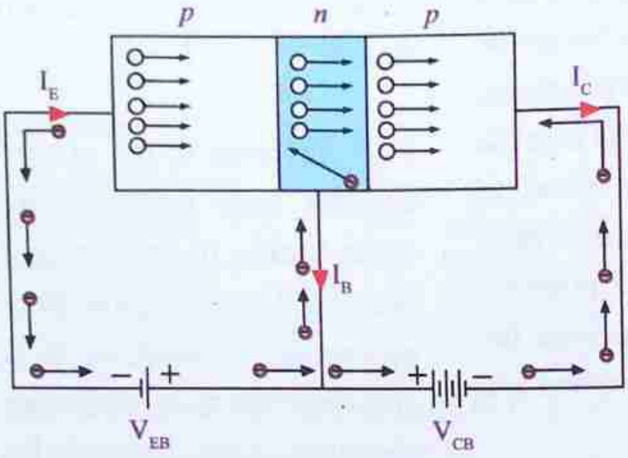


Figure 5.37: *Mode of action of p-n-p transistors*

The forward bias causes the holes in the *p*-type emitter to flow towards the base. This constitutes the emitter current, I_E. As these holes cross into *n*-type base, they tend to combine with the electrons. Since the base is lightly doped and very thin, only few holes (about 5%) combine with the electrons to constitute the base current I_B.

The remaining holes, cross into the collector region to constitute collector current, I_C. Again,

$$I_E = I_B + I_C$$

Uses of transistors

Transistors are said to be the basic elements of modern electronics. They are used in all electronic devices including calculators, televisions, radios and computers. Transistors are used in switching circuits, amplifier circuits, oscillator circuits, current source circuits, voltage regulator circuits, power supply circuits, digital logic integrated circuits and in any circuit that uses small signals to control large currents.

Information signals

Information is usually transmitted in electronic devices in form of signals. Information signals are divided into two broad classes, namely *analogue* and *digital* signals. The process of encoding information into analogue signals is different from encoding information into digital signals. Consequently, analogue and digital signals are decoded differently. All operations that can be performed on an analogue signal, such as amplification, filtering, limiting and others, can also be done in the digital domain.

Analogue signals

Analogue signals are electrical signals that convey or store information by means of variation in a continuous waveform. The first electronic devices that were invented and produced in large numbers are analogue devices. Figure 5.38 shows an example of analogue signal.

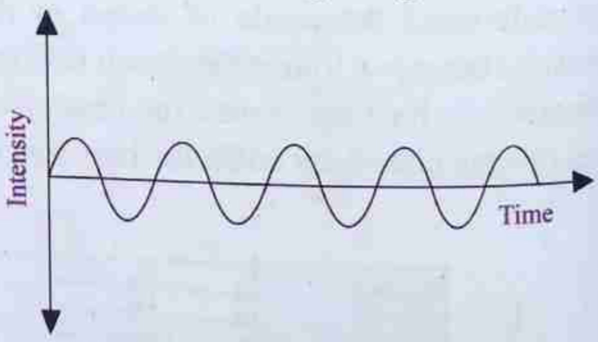


Figure 5.38: *Analogue signal waveform*

The signal is basically an electric current or voltage that is produced due to a physical change. Examples of physical change include sound volume, light intensity, temperature level and pressure level. The physical change is converted to electrical current by a device known as *transducer*.

A transducer is a device that converts an input signal of one form into an output signal of another form.

Analogue signals may take any value from a given range, and each unique signal value represents different information. Any change in the signal is meaningful, and each level of the signal represents a different level of the phenomenon that it represents. For example, if the signal is being used to represent temperature, with one volt representing one degree Celsius, then 10 volts would represent 10 degrees celsius, and 10.1 volts would represent 10.1 degrees celsius.

In an analogue sound recording, the variation in pressure of a sound striking a microphone creates a corresponding variation in the current passing through it or voltage across it. An increase in the volume of the sound causes the fluctuation of the current or voltage to increase proportionately while keeping the same waveform or shape. Mechanical, pneumatic, hydraulic and other systems may also use analogue signals.

Digital signals

Digital signals, unlike analogue signals, are non-continuous electrical signals. They change in steps. They convey information in pulses or digits of two discrete levels. This means that the value of each pulse is constant and moving from one digit to the next is an abrupt change. Figure 5.39 shows digital signal waveform.

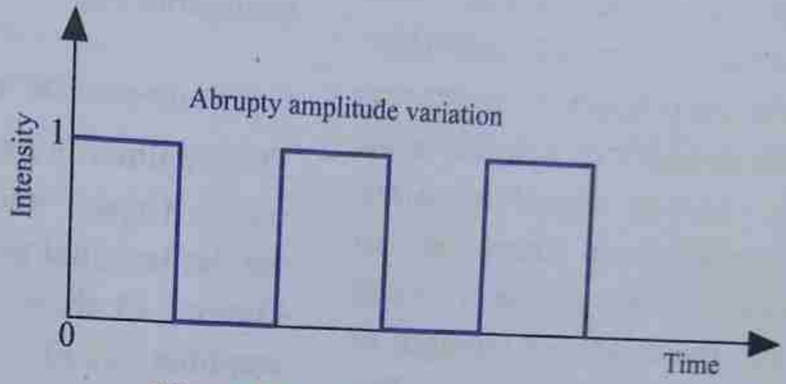


Figure 5.39: *Digital signal waveform*

Digital signals have only two amplitude levels, usually called nodes. This means the values can only be given in one of two ways. The values may be specified, for example, as 1 or 0, TRUE or FALSE, or, HIGH or LOW. Digital signals are often derived from analogue signals.

The main advantage of digital signals over analogue signals is that the signal level need not to be precise. It can be approximated within a fixed number of digits or bits, as illustrated in Figure 5.40.

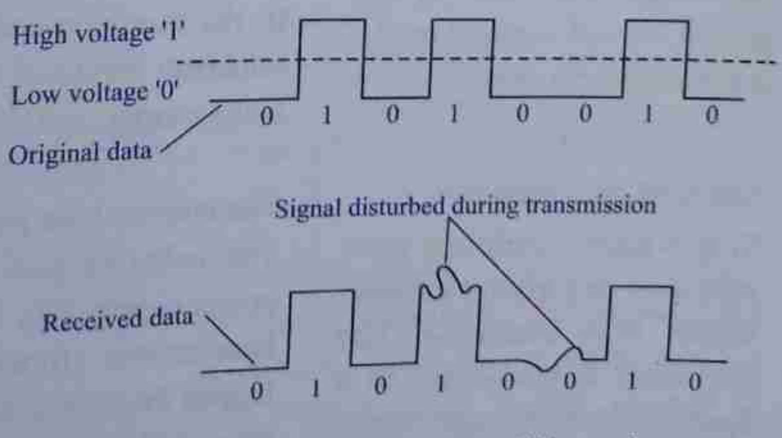


Figure 5:40: *Distorted digital image can still be read correctly*

Electronic amplifiers

Analogue signal may take different values. However, very small signals are practically undesirable. If the signals are very small (weak) a specific device is used to increase the strength of the signals. This device is called *amplifier,* and the process of increasing the strength of a weak input or output signal is called *amplification*. Therefore, an amplifier is a device or circuit that increases the strength of the weak input signals. The output signal of an amplifier is related to the input signal by a transfer factor known as the amplifier gain. There are many types of amplifiers commonly used in radio and television transmitters and receivers, stereo systems, microcomputers and musical instruments. These can usually be simple circuits, for example *single-stage,* or complex circuits, for example *multistage amplifiers*. In this section, we will consider single-stage amplifiers only.

Single-stage amplifiers

Single-stage amplifiers have only one amplifying element which is a transistor. The transistor is connected to a load resistor through which a collector current flows. The value of the load resistor together with the transconductance value affects the amplifier's voltage gain. Single-stage amplifiers may be grouped according to the connection or configuration of the transistor. Therefore, there are three types of single stage amplifier connections (configurations). These are common-emitter (CE) amplifier, common-collector (CC) amplifier and common-base (CB) amplifier configurations.

Common-emitter amplifier configuration

The amplifier is named so because both the signal source (input) and the load (output) share the emitter terminal as the common connection point. Figure 5.41 shows a single stage common-emitter amplifier circuit.

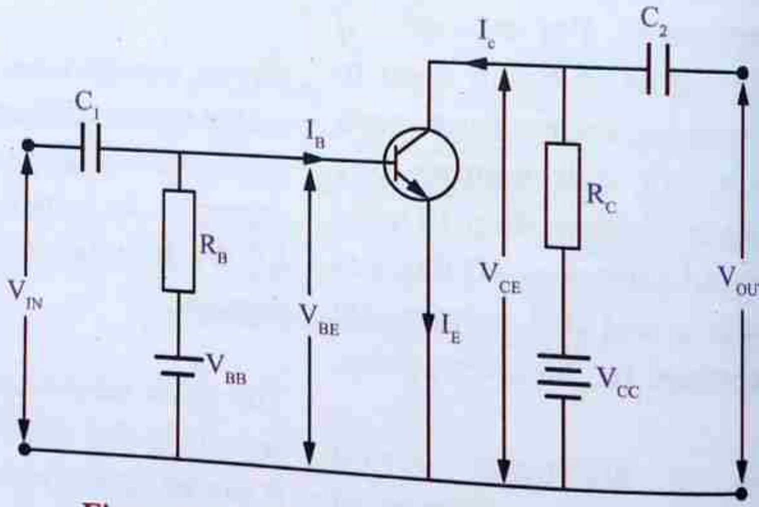


Figure 5.41: *Common-emitter amplifier circuit*

In the common-emitter amplifier, the base of the transistor serves as the input terminal, the collector is the output, and the emitter is common to both.

The emitter-base junction is forward biased by V_{BB}. The collector-base junction is reverse biased by power supply V_{CC}. The input signal, V_{IN} is fed to the base-emitter circuit and the output signal, V_{OUT} is tapped from the collector terminal with respect to the grounded emitter terminal. C_1 and C_2 are coupling capacitors to provide d.c isolation at the input and output of the amplifier respectively. The emitter current is the sum of base current and collector current, $I_E = I_B + I_C$.

Current gain

In common-emitter configuration, the input current is base current, I_B and the output current is collector current, I_C. The ratio of collector current, I_C to base current, I_B is defined as the current gain, β of common-emitter transistor.

That is,

$$\beta = \frac{I_C}{I_B}, \text{ for } V_{CE} = \text{constant}$$

Note that the value of base current is quite low compared to the collector current ($I_B << I_C$), therefore the current gain, β in a common-emitter transistor is quite high and it ranges from 20 to 500.

Example

In a common emitter circuit, calculate the emitter current, I_E of a transistor for which current gain, β is 50 and base current, I_B is 20 μA.

Solution

The current, I_E is obtained from the relation $I_E = I_B + I_C$.

But I_C can be obtained from the relation $\beta = \frac{I_C}{I_B}$.

Therefore,

$$I_C = \beta I_B$$

$$= 50 \times 20 \times 10^{-6} \text{A}$$

$$= 1 \times 10^{-3} \text{A} = 1 \text{ mA}$$

Using the relation,

$$I_E = I_C + I_B$$

$$= 1 \text{ mA} + 0.02 \text{ mA}$$

$$= 1.02 \text{ mA}$$

$$\therefore I_E = 1.02 \text{ mA}$$

Common-collector amplifier configuration

In this type of amplifier, the base of the transistor serves as the input terminal, the emitter as the output, while the collector is common to both. Figure 5.42 shows a common-collector amplifier circuit.

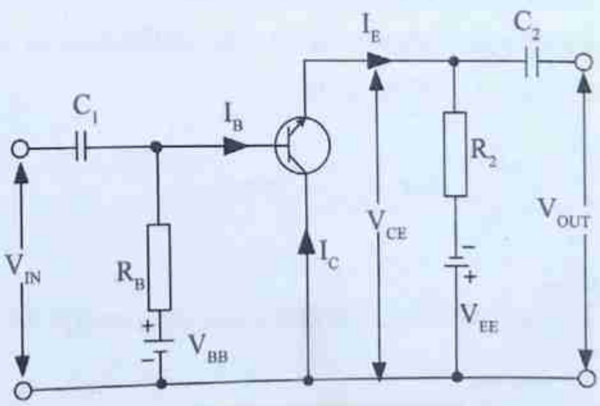


Figure 5.42: *Common-collector amplifier circuit*

The input signal is fed to the base-collector circuit while the output signal is tapped from the emitter terminal with respect to the ground. C_1 and C_2 are coupling capacitors to provide d.c isolation at the input and output of the amplifier respectively.

Common-base amplifier configuration

In the common-base amplifier, the emitter serves as the input terminal, the collector as the output, while the base is common to both. Figure 5.43 shows a common-base amplifier circuit.

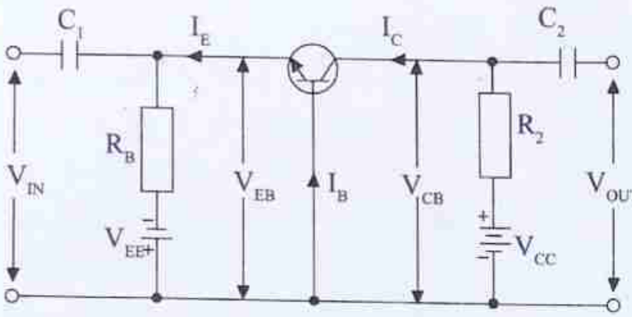


Figure 5.43: *Common-base amplifier circuit*

The emitter-base junction is forward-biased by the power supply V_{EE} whereas the collector-base junction is reverse-biased by V_{CC}. The input signal is fed to the emitter-base circuit while the output signal is tapped from the collector-base circuit. Current gain,

$$\beta = \frac{I_C}{I_E}, \text{ for } V_{CB} = \text{constant.}$$

The value of α lies between 0.95 and 0.99.

Activity 5.1

Aim: To demonstrate the working principle of a transistor.

Materials: Resistors (100 kΩ – 470 kΩ), *n-p-n* transistor, connecting wires, micro-ammeter (μA), milli-ammeter (mA), d.c power supply or batteries enough to provide 6 V and 3 V.

Procedure

1. Set up the circuit as shown in Figure 5.44.

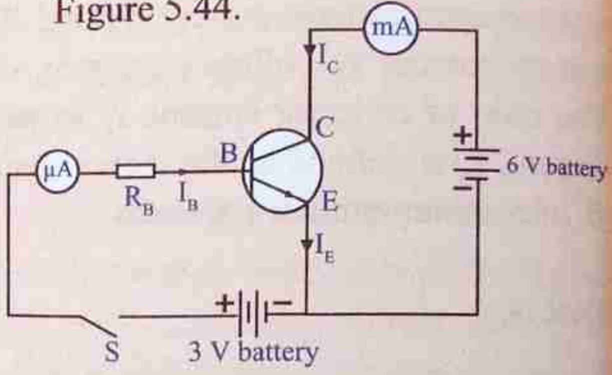


Figure 5.44: *Common – emitter circuit*

2. Close the switch (S) to provide the supply to the base.
3. By using different resistors ranging from 100 kΩ to 470 kΩ obtain the respective readings of I_B and I_C.

Questions

(a) Find the emitter current I_E.

(b) Suppose 20 Ω is used in step 3. What will happen to I_B?

Task

1. Visit a nearby electronic workshop and observe how amplifiers work. Make a report and present it to your teacher.
2. With the help of your teacher, design a single-stage amplifier circuit. Use it to amplify current.

Exercise 5.2

1. (a) What is a transistor?
   (b) Transistors are classified as bipolar junction and field effect transistors. What is the criteria on which this classification is based ?
2. The bipolar junction transistor is of two categories: *p-n-p* and *n-p-n*.
   (a) What is the difference between them?
   (b) Draw the symbol for each transistor category.
3. Explain why the base of the transistor is made very thin.
4. Explain the two common uses of a transistor.
5. Explain the main characteristic of a single stage amplifier. State the function of each component of this amplifier.

Chapter summary

1. Electronics is a branch of physics that deals with the emission, behaviour and effects of flow of electrons in semiconductors, vacuum or gas and devices.
2. Circuit elements that do not have the ability to amplify signals are called passive elements. These include resistors, inductors and capacitors.
3. Active elements such as transistors and ICs can amplify a signal.
4. All semiconductors have electrical conductivity that is intermediate between that of insulators and conductors. The electrical conductivity of semiconductors increases with a increase in temperature.
5. An intrinsic semiconductor is a semiconductor which is very pure such that any impurities in it do not significantly affect its electric behaviour.
6. The electrical conductivity of semiconductors can be improved by introducing a controlled amount of impurity in it.
7. Donor atoms add electrons to a semiconductor producing an *n*-type semiconductor while acceptors add electron vacancies called holes producing a *p*-type semiconductor.
8. A *p-n* junction is formed where *n*-type and *p*-type regions are placed in contact to each other. The *p-n* junction can only pass current in one direction and is the basis of rectifiers used to convert a.c to d.c.
9. A diode is an electrical component that allows electric current to flow through it in one direction with far greater ease than in the other. A diode consists of a single *p-n* junction with conductive contacts and wire leads connected to each region.

10. Diodes are used to rectify a.c. into d.c in two ways, namely full-wave rectification and half-wave rectification.
11. A transistor is made of two sections of one semiconductor type separated by a section of the other semiconductor type. The centre region is called the base of the transistor while the other two regions are called the emitter and the collector.
12. There are two types of bipolar junction transistors. These are *n-p-n* and *p-n-p* transistors.
13. Transistors are used as switches or amplifiers. Single-stage amplifiers make use of a single transistor to amplify electrical signals.
14. A single stage amplifier can be made in three configurations, namely, common-emitter (CE), common-collector (CC) and common-base (CB).

Revision exercise 5

Choose the correct answer for items 1-5.

1. An intrinsic semiconductor produces ________ semiconductor when a pentavalent atom is added to it.
   (a) extrinsic type
   (b) *p*-type
   (c) *n*-type
   (d) hybrid type
2. Which biasing results into large diode current?
   (a) Forward bias.
   (b) Reverse bias.
   (c) Inverse bias.
   (d) Knee bias.
3. When a *p-n* junction is connected to a battery in such a way that *p*-side is connected to negative terminal of the battery and positive terminal to *n*-side, the connection is known as:
   (a) Forward bias.
   (b) Feverse bias.
   (c) Zero bias.
   (d) Knee bias.
4. What is the main characteristic of electrons in the conduction band?
   (a) Lose their charge easily.
   (b) Jump to the tip of the crystal.
   (c) Have a higher energy than the valence band.
   (d) Have lower energy than that of electrons in the valence band.
5. The general characteristics of a common-base amplifier are:
   (a) High voltage gain, low current gain, high power gain and low input resistance.
   (b) High voltage gain, low current gain and low input resistance.
   (c) Low voltage gain, high current gain, very high-power gain and low input resistance.
   (d) High voltage gain, high current gain and high input resistance.

6. Describe the effect of temperature on conductors, semiconductors and insulators.
7. Explain the difference between conductors, semiconductors and insulators using the energy band theory.
8. (a) What is doping?
   (b) Explain the mechanism of doping.
9. (a) What is a *p-n* junction?
   (b) Explain the mode of action of a *p-n* junction.
10. (a) List down any two types of diodes.
    (b) Draw a circuit diagram to show half-wave and full-wave rectification.
11. (a) What is a transistor?
    (b) Draw a diagram of a *p-n-p* transistor and explain its mode of action.
    (c) List down two applications of transistors.
12. Figure 5.45 shows some circuit diagrams for single-stage amplifiers.

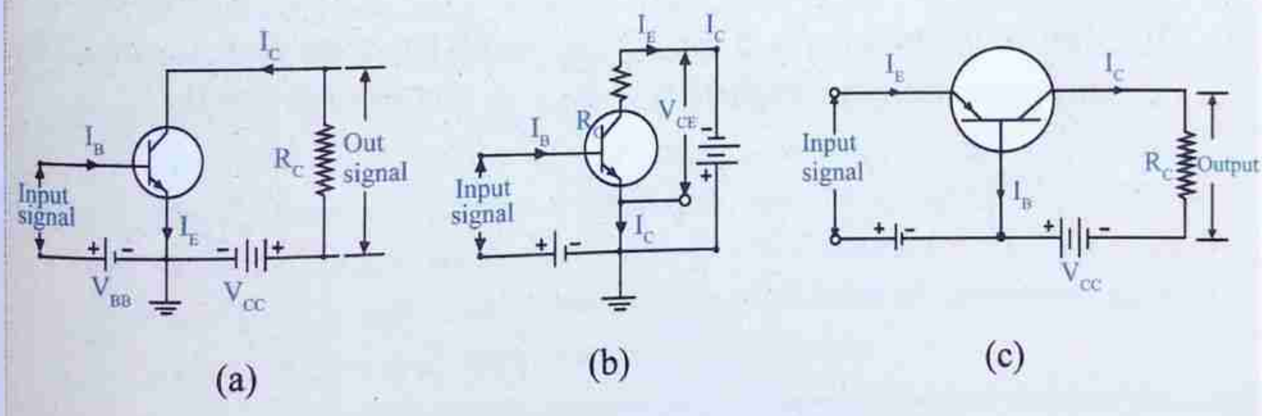


Figure 5.45

Identify the type of amplifier configuration represented by each circuit diagram in Figure 5.45. Give reasons for your answer.

13. Figure 5.46, shows a simple d.c power supply circuit.

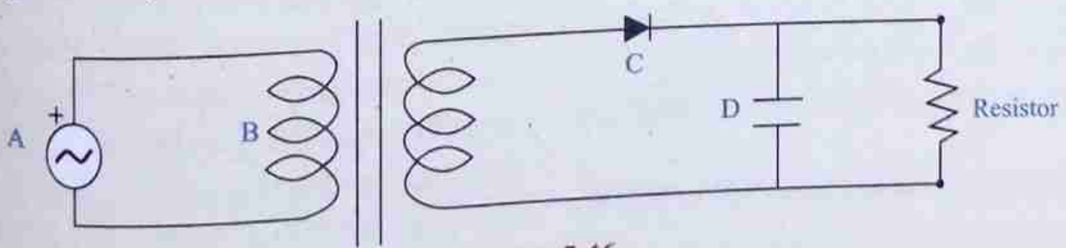


Figure 5.46

(a) What does the circuit symbols A, B, C, and D stand for?
(b) What are the functions of components A, B, C and D in the circuit?

14. (a) Define the term semiconductor.
    (b) How does an intrinsic semiconductor differ from extrinsic semiconductor?
15. (a) With the aid of diagrams, differentiate a Zener diode from a LED.
    (b) Explain the main characteristics of conduction band and valence band.
16. (a) Describe the mechanism of doping an intrinsic semiconductor to obtain *p*-type semiconductor.
    (b) Most of the transistors in use are *n*-*p*-*n* transistors. Explain.
17. The collector current, I_C in a common-emitter amplifier is directly proportional to the base current I_B.
    (a) What does this statement mean?
    (b) Write down the mathematical expression for the statement.
    (c) If the current gain, β of this amplifier is 200, calculate the collector current given that the base current I_B is 2 μA.
18. (a) Write an expression that relates I_E, I_C and I_B for a CE amplifier configuration.
    (b) If $I_C = 3.25$ mA and $I_E = 3.3$ mA f; calculate I_B.
    (c) What is the current gain of the transistor in part (b)?

Chapter Six

Elementary astronomy

Introduction

Watching the sky after sunset is really fascinating. You would first notice one or more bright dots shining in the sky. Soon you will realise that, the number of bright dots is increasing until you cannot count them anymore. The entire sky is filled with countless objects such as stars twinkling with varying brightness depending on how far they are from the Earth. Among these bright objects, you may also be able to see the Moon (Earth's natural satellite) and the naked-eye planets. You may also see artificial satellites that have been launched for various purposes including communication and space surveys. The science of observing the physical universe is called astronomy. In this chapter, you will learn about the concept of astronomy, constellations, the solar system, gravitational force as well as the Earth and its Moon. The competencies developed will enable you to use simple skills of science and available local materials to explore the sky at night and estimate some astronomical parameters such as time for ocean tides and relative positions of planets.

Concept of astronomy

Astronomy involves the study of the physical universe and the objects in it that are beyond the Earth's atmosphere. The universe encompasses space, time, energy and matter. The objects (matter) of most concern in *astronomy* includes galaxies, stars, planets, moons, asteroids and comets. The word astronomy is derived from two Greek words *astron* and *nomos* for "star" and "laws or cultures", respectively. Therefore, the literal meaning of the word astronomy is "law of the stars". Figure 6.1 shows some objects in the universe. Currently, the size of the observable universe is said to be about 46 billion light-years (that is, light with its speed, 3×10^8 m s $^{-1}$ takes about 46 billion years to travel from one edge to the opposite edge of the observable universe).

Astronomy is a science that deals with the study of objects and phenomena beyond the Earth's atmosphere.

Figure 6.1: *Some objects in the universe*

Astronomers

People who study astronomy are called *astronomers*. Astronomers study objects that compose the physical universe such as stars, planets, satellites, nebulae, galaxies as well as distribution of matter and energy in space and time. Astronomers can be categorized as professional astronomers and amateur astronomers. Due to the complexity of the universe, astronomy has become a multidisciplinary field that cuts across all disciplines. Astronomy therefore, applies principles, theories and laws of all natural Sciences, Technology, Engineering and Mathematics (STEM) as well as social sciences. This makes it one among the most interesting fields of study for all generations.

The importance of astronomy

Astronomy is one of the oldest fields of natural science. Early astronomy was limited to observations of regular patterns of motions of visible celestial bodies including stars, planets, natural satellites, and transit objects such as comets, meteoroids and meteors, using naked-eyes. Despite the limited technology, astronomy had several applications in the old days. Some of the uses of astronomy in the ancient period are hereby discussed.

1. It was the earliest method of measuring time. A day was a time duration between sunrise and sunset while a month was derived from phases of the Moon. A year was derived from the changing positions of sunrise.
2. It was used to develop calendars that made it possible to predict the seasons due to Earth's revolution. The seasons were very important in agriculture as they dictated the planting and harvesting time. Star constellations were used to determine the different seasons for growing crops.
3. It was used in both land and sea navigation, based on the knowledge of the positions of the Sun during the day and positions of other stars at night.

The development of science and technology has greatly boosted the astronomers' abilities to study the universe.

Consequently, modern astronomy is credited with improved knowledge of the universe and more advanced applications such as:

1. Knowledge of the origin, composition and evolution of the universe.
2. Improved navigation on land, sea and air. For example, through the Global Position System (GPS).
3. Better knowledge of natural phenomena such as gravitational pull.
4. A frontier for multidisciplinary and multinational collaboration in research and exploration of nature for social economic benefits. The international space station is a good example of this.
5. Development of Astro-tourism including space tourism. For example, tourism at Mbozi meteorite (Figure 6.2), Ndolezi village, Mbozi district in Songwe region, Southern Tanzania. The Mbozi meteorite is estimated to weigh 12 tones with dimensions of 1.22 metres in height, 1.63 metres in width and 3.3 metres in length. This is among the 10 heaviest meteorites in the world and the second heaviest meteorite in Africa.

Figure 6.2: *Mbozi meteorite*

Further more the scientific and technological development in astronomy has contributed to other fields such as medicine, defense and energy .

Galaxies

A galaxy is a huge collection of star systems (stars and their solar systems), interstellar nebulae (gas and dust) and dark matter held together by gravitational force. Currently, there are about 100 to 200 billion galaxies in the observable universe. Galaxies are classfied based on their shape as spiral, elliptical or irregular. The Sun as the nearest star to the Earth forms a solar system which belongs to a galaxy called the Milky Way.

The Milky Way Galaxy is a spiral galaxy (Figure 6.3 (a)). It contains both old and young stars. It has a diameter of about 100,000 light-years. Nearly all of the stars visible in the night sky belong to Milky Way Galaxy. The hazy band of light that can be seen in a dark clear night sky as shown in Figure 6.3 (b) is the Milky Way Galaxy.

(a) Size the of Milky Way Galaxy

(b) Milky Way Galaxy as viewed in a dark clear night sky

Figure 6.3: *Milky Way Galaxy*

Exercise 6.1

1. Explain the term astronomy and state its importance.
2. Briefly explain the composition of the universe.
3. (a) Mention and describe briefly the galaxy to which our solar system belongs.
   (b) What is the shape of the galaxy described in part (a)?

Stars

A star is a large celestial body made up of hot luminous ionized gases known as plasma, held together by gravity. Plasma is an ionised gas that contains certain proportion of electrons that are free, not bound to an atom or molecule. The main gases composing a star are hydrogen and helium. Stars shine as they burn hydrogen into helium in their cores, through a process of nuclear fusion. The nuclear fusion process generates enormous amount of energy which is radiated away in the form of light.The twinkling of stars is due to the light variation as it undergoes refraction when travelling through atmospheric layers of different refractive indices.

The Sun is the closest star to the Earth at a mean distance of about 150 million kilometres. This distance is known as the Astronomical Unit, abbreviated as AU. It is used to measure distances across the solar system.

Note: 1 AU = 150 million kilometres.

Star formation

Stars are formed from a collection of interstellar nebula (gas and dust) that has collapsed down from a giant molecular cloud (star forming region). A young star that gathers nebula from its parent molecular cloud before it collapses to form a new star is called a protostar (Figure 6.4). There are many stars in the universe, but of our particular interest is the Sun which gives energy to our solar system.

Figure 6.4: *Protostar showing a new star being born*

Constellations

Constellations are small groups of bright stars that form patterns in the sky that resemble the familiar animals or objects on the Earth. Ancient people especially Greeks named constellations after mythological characters, heroes, animals and other things on Earth.

Currently, there are about 88 officially known constellations. Various constellations are visible during a particular period of the year. Some examples of the known constellations are shown in Figure 6.5.

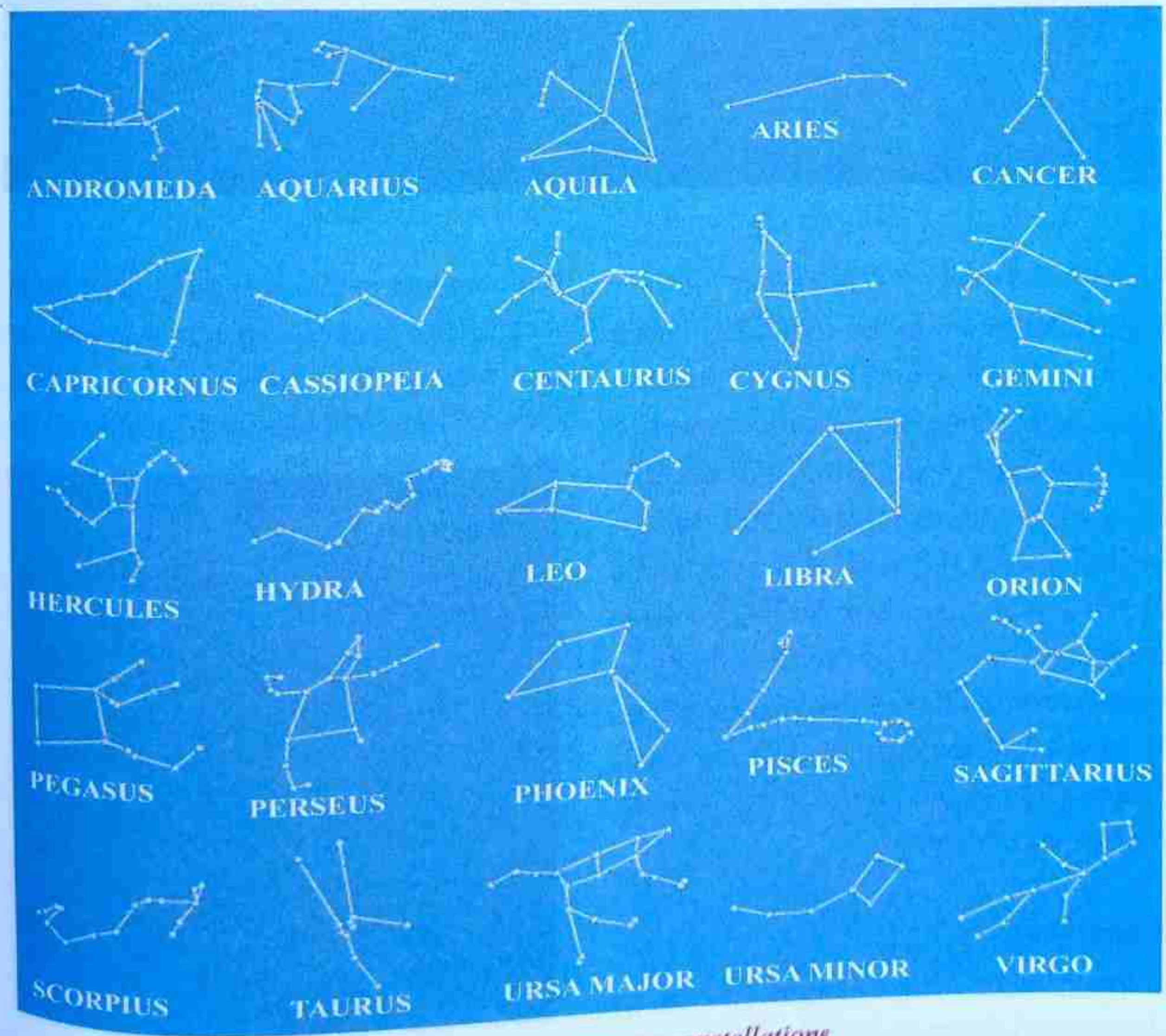
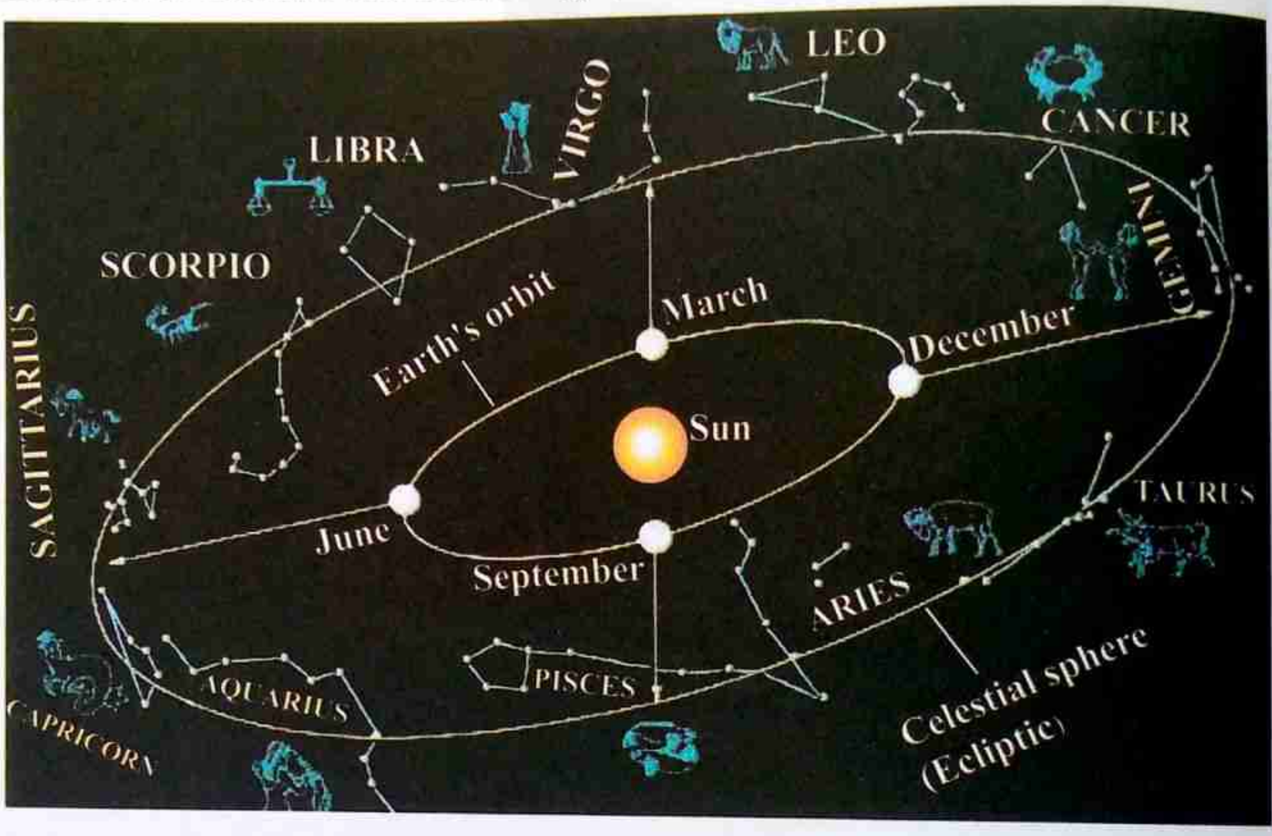


Figure 6.5: *Some common constellations*

Some constellations form a set of 12 constellations in the same plane as the Earth's orbital plane around the Sun as shown in Figure 6.6. This set is known as *Zodiac* (It should not be confused with zodiacal light).



Identification of constellations

Constellations may be identified based on their shape (morphology), position (location) in the sky or with respect to a neighbour constellation, the number of stars in the pattern, and the time of visibility over the year. However, some characteristics may look similar for some constellations. Some few examples of common constellations that can sometimes be seen during the night sky include: Canis Major, Ursa Major, Puppis, Leo, Germini, Cancer and Orion.

Canis Major

Canis Major constellation is easily determined since it includes Sirius (the brightest star in the sky). In Latin its name means 'the greater dog' which represents the bigger dog following the hunter Orion. Figure 6.7 shows the Canis Major.

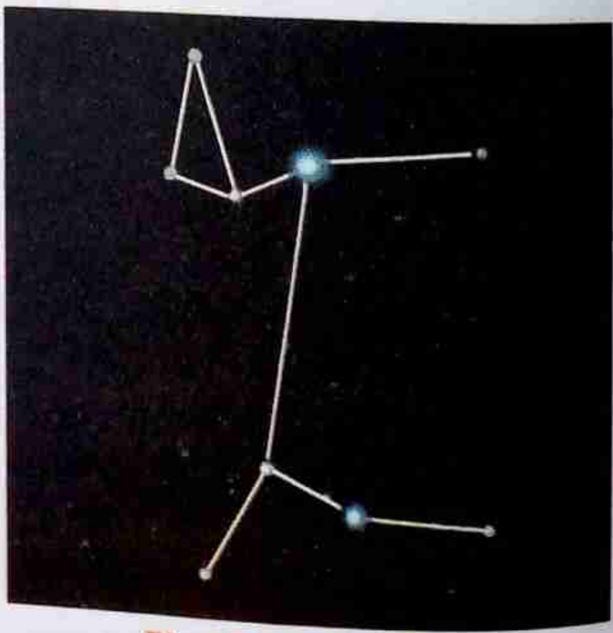


Figure 6.7: *Canis Major*

Ursa Major

Ursa Major looks like a great bear, making it to be called the *Great Bear.* It is right opposite the Orion. Ursa Major is the third brightest constellation in the night sky. It has seven prominent stars. Figure 6.8 shows the Ursa Major as seen in the night sky.

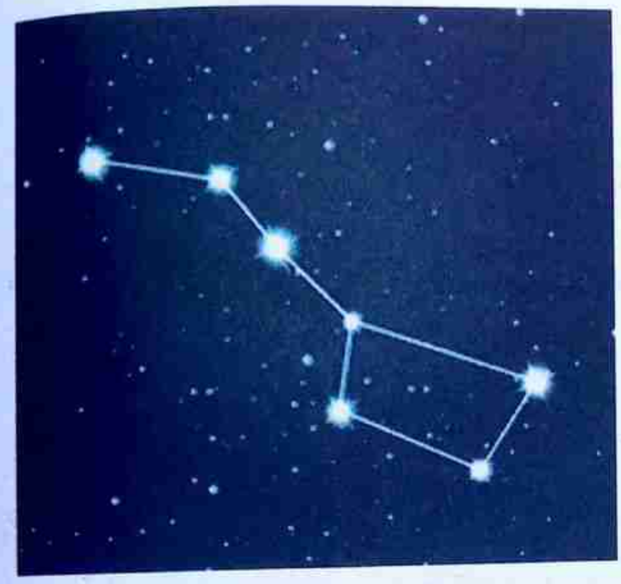


Figure 6.8: *Ursa Major*

Puppis

It is a pattern of nine stars right next to the Canis Major. The brightest star in this constellation is known as *Zeta Puppis*. Figure 6.9 shows a Puppis constellation.

Figure 6.9: *Puppis*

Leo

It is a lion-like in shape constellation. Leo is one of the Zodiac constellations with most of the stars visible. Figure 6.10 shows Leo constellation.

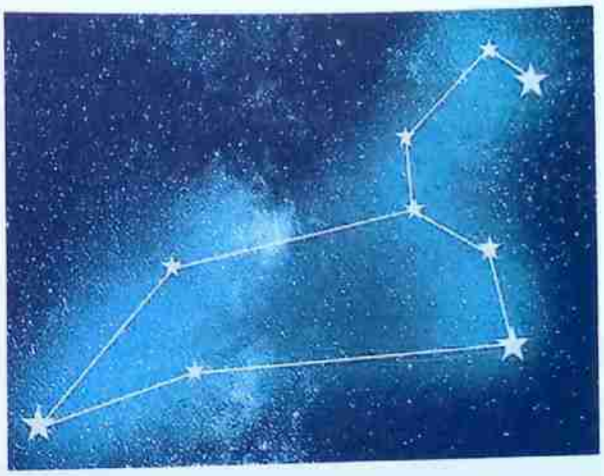


Figure 6.10: *Leo*

Gemini

It is one of the constellations in the zodiac. Its name originates from Latin word *Gemini* which means twins as seen in Figure 6.11. It is located northeast of the Orion constellation.

Figure 6.11: *Gemini*

Cancer

It is a crab like constellation that looks like an upside-down letter Y and the faintest of the zodiac constellations. Cancer is located between Leo, the lion, and Gemini, the twins. Figure 6.12 shows the Cancer constellation.

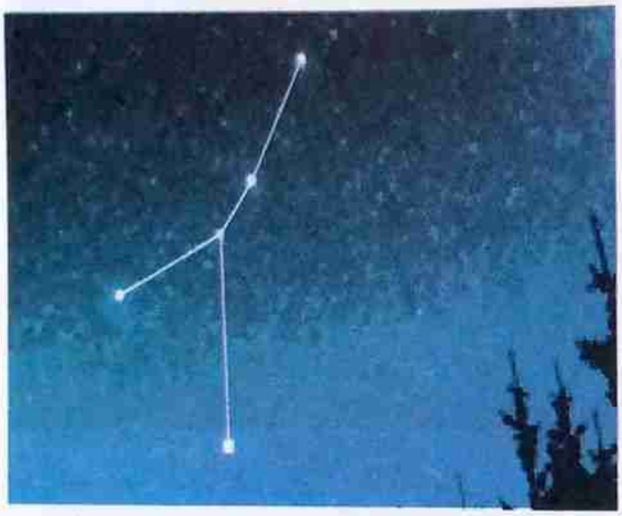


Figure 6.12: *Cancer*

Orion

It looks like a hunter with a bow and arrow. Orion is visible in the evening sky during winter in the Northern Hemisphere from January to March and during summer in the Southern Hemisphere. Figure 6.13 shows the Orion constellation.

Figure 6.13: *Orion*

Activity 6.1

Aim: To demonstrate a modal of a constellation.

Materials: Batteries, connecting wires, light emitting diodes (LEDs) of different colours indicating different star brightness, pencil and pen, soft board for drawing the constellation pattern

Procedure

1. Select the constellation that you want to demonstrate.
2. Sketch its diagram on a soft board using any drawing material.
3. Connect the given LEDs using connecting wires.
4. Complete the pattern of your LEDs that represent stars by connecting to a source of electricity.
5. Observe the pattern as the LEDs twinkle.

Questions

(a) Which type of constellation did you select?

(b) What are the names of the stars found in your selected constellation?

(c) Did your pattern resemble the natural constellation?

Uses of constellations

The following are some uses of constellations:

1. **Religious**

In early days, people thought that the gods and goddesses lived in the heavens and that the gods created the constellations. Many cultures believed that the positions of the stars were their god's way of telling stories. Indeed, the Greeks named the constellations after their mythological heroes and legends. For example, to the ancient Greeks, Orion was a great hunter. He was the son of Neptune (god of the sea).

2. **Agriculture**

Before the existence of proper calendars, people had no way of determining when to sow or harvest except by observing changing positions of celestial objects such as stars. Constellations made the patterns of the stars easy to remember. The ancient people knew, for example, when the constellation Orion started to be fully visible, winter was about to start. The constellations assisted farmers to plan ahead.

3. **Navigation**

It is fairly easy to spot Polaris (The North Star) once you have found Ursa Minor (Little Dipper constellation). One can figure out his or her latitude just by looking at how high Polaris appears in the night sky. This enabled marine vessels to travel across the globe in the intended direction. Polaris was used in the discovery of America, the spread of European culture and civilisation as we know it today.

4. **Nomads**

Nomads used constellations in travelling with their herds from one place to another, searching for pasture.

5. **Curiosity of studying nature**

Constellations as other celestial objects, trigger humankind to explore more about the universe.

Task 6.1

Observe a clear night sky for at least two days in order to identify some constellations visible at your respective location.

1. In groups share the experience of the observed constellations and draw any five constellations on manila paper.
2. Present the drawn constellations in class for a collective discussion.

Exercise 6.2

1. (a) How are stars formed?
   (b) Why do stars shine?
2. Briefly explain the meaning of constellation. Give two examples of constellations.
3. Distinguish between zodiac and zodiacal light.

Sun

The Sun is the star at the centre of our solar system whose mass is approximately 2.0×10^{30} kg. The Sun emits radiant energy which reaches the Earth's surface in forms of light. The radiant energy comes from continous fusion in the interior of the Sun.

Structure of the Sun

The structure of the Sun comprises of a core, radiative zone, convective zone, photosphere, chromosphere, transitional zone and corona as shown in Figure 6.14.

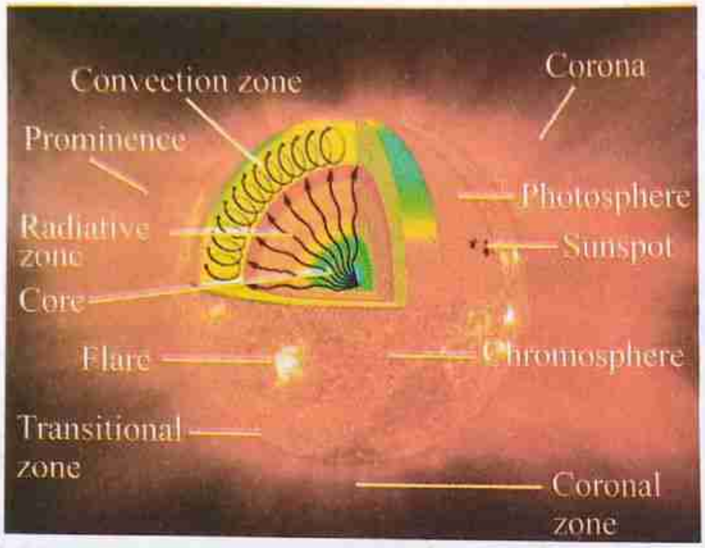


Figure 6.14: *Interior structure of the Sun*

Core:

A core is the innermost part of the Sun. It is the densest and hottest region having temperature ranging from 10-15 million Kelvin. Nuclear fusion reactions create energy within the core resulting to gamma rays and chargeless particles called neutrinos. The core size is about 200 000 km.

Radiative zone:

It is found outside the core where energy is transported by radiation. The radiative zone size is about 300 000 km.

Convection zone:

It is the zone after the radiative zone where materials carrying energy are transferred in form of convection. The convection zone has the same size as the core, at 200 000 km.

Photosphere:

It is the outer part of the Sun that emits electromagnetic waves, including visible light. The spectrum of the emitted radiation helps to study the characteristics of the star such as temperature and chemical composition. The photosphere contains sunspots which are cooler and darker points. They are cooler because strong magnetic field loops around them prevent the heat from reaching the surface. The photosphere size is about 500 km.

Chromosphere:

It is the reddish part after the photosphere which is rich in hydrogen gas. Using special filters, it can be seen during a total solar eclipse as a red haze around the sun. It can also be seen during solar flares, that is fire like charged particles that emerge from sunspots in the photosphere. The chromosphere size is about 1500 km.

Transitional zone:

It is a narrow layer between chromosphere and corona. Its size is about 8 500 km. Its temperature rises abruptly up to several million kelvins.

Corona:

The outermost part of the Sun. It is visible during a total solar eclipse. Sometime, it can be viewed by using coronagraph. It consist of immense clouds of glowing gas called prominences that erupt from the upper chromosphere and shoot into the corona. Corona is the source of solar winds which sometimes shoot onto the Earth's atmosphere and may interfere with power grids and telecommunication networks.

The solar system

The solar system is made up of the Sun and all the objects bound to it by gravity. The objects revolve around the Sun in fixed paths known as orbits. The objects that orbit the Sun include the eight planets, their known moons and billions of small bodies including asteroids, comets and interplanetary dust. Figure 6.15 shows some of the objects of the solar system.

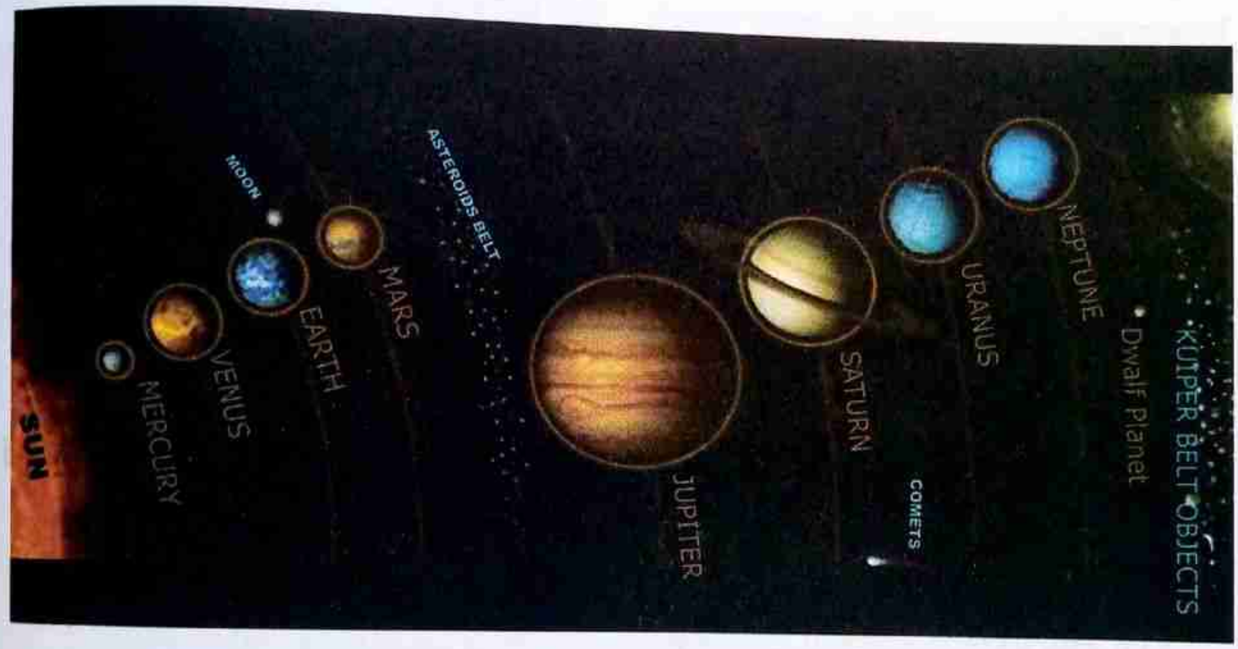


Figure 6.15: *Solar system*

Theories of the solar system

Description of the structure of the solar system has involved different theories that explain how the Sun, Earth and other celestial objects are arranged and how they move. The basic theories are the geocentric and heliocentric theories.

Geocentric theory

The most predominant theory of the structure of the universe in the ancient world is the geocentric theory. This theory assumed that the Earth is at the centre of the universe and other celestial objects rotate around the Earth. Figure 6.16 shows a solar system according to the geocentric theory.

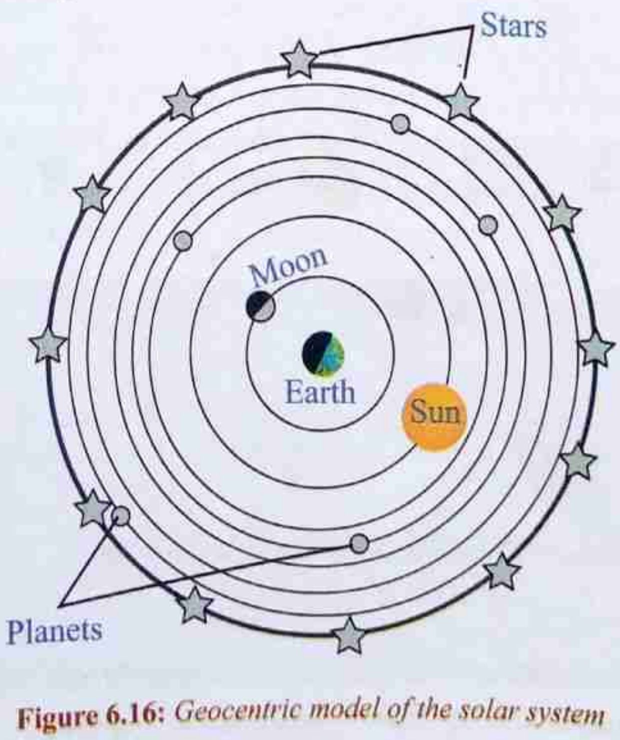


Figure 6.16: *Geocentric model of the solar system*

Heliocentric theory

The heliocentric theory holds that the Sun is at the centre of the universe and other celestial objects revolve around it. Figure 6.8 illustrates the idea of the solar system according to the heliocentric theory.

This theory opposes the geocentric theory and is able to explain many of the observations made by astronomers.

These observations include:

(a) Rotation of the Earth about its own axis and its revolution around the Sun.

(b) Variations in brightness of Venus and Mars.

(c) Clarification for the positions of Venus and Mars with respect to the Sun.

(d) Logical estimation of the size of the Sun with respect to that of the Earth.

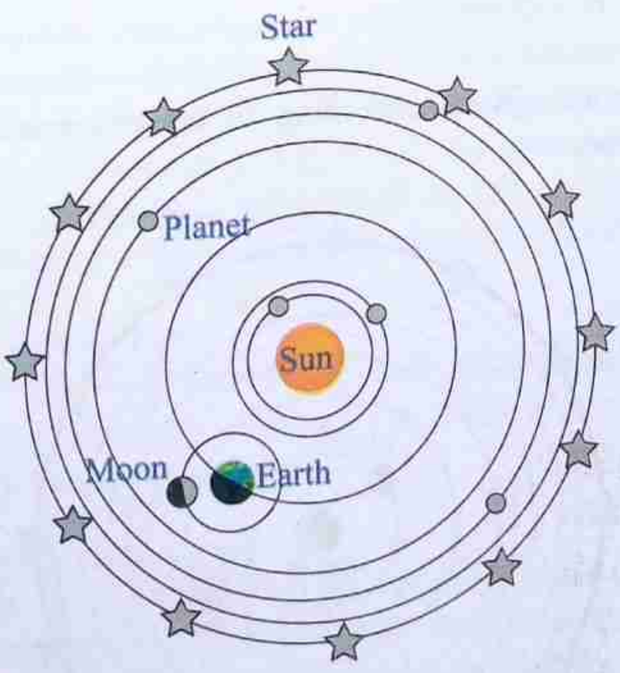


Figure 6.17: *Geocentric model of the solar system*

The heliocentric theory is currently the basic theory that is adopted to explain the solar system. However, the theory has a few shortfalls, which include:

1. The Sun was thought to be the centre of the universe but it is only the centre of the solar system.
2. Orbits of the planets were thought to have circular shapes, but nowadays they are known to have elliptical shapes.
3. The relative positions of the stars seemed to remain the same despite the Earth's changing view points as it moved around the Sun.

Activity 6.2

Aim: To investigate the relationship between distance, size and visibility of an object.

Materials: Three white balls of the same size about 40 cm in diameter or any other objects but of the same size and colour, tape measure, a place that can provide a distance of at least 100 m, for example a playground

Procedure

1. Measure the distance of 10 m, 50 m and 100 m from one corner of the playground along the edge.
2. Place the first ball at the distance of 10 m.
3. Place the second ball at the distance of 40 m from the first ball in the same line as the first ball.
4. Place the last ball at the distance of 50 m from the second ball. Ensure

all balls are in the same straight line as illustrated in Figure 6.19.

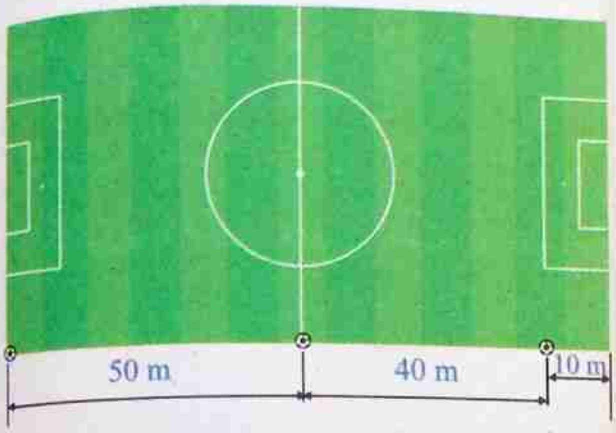


Figure 6.18

5. Stand at one end of the playground.
6. Using your naked eyes, observe and compare the ball colour visibility, size and distance.

Questions

(a) Which ball looks brighter (more visible) than the rest?
(b) Which ball looks fainter as compared to the rest?
(c) What do you conclude about the relationship between the distance of an object and its colour visibility?
(d) How did the ball size apparently vary with distance?
(e) From these observations, explain why the Sun looks brighter and bigger than any other observable star?

Planets

A planet is a celestial object which orbits a star. Currently, there are eight planets orbiting the Sun. Starting with the closest planet from the Sun, these are Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus and Neptune. Among these, Mercury, Venus, Mars, Jupiter and Saturn can be directly seen by our eyes. These planets are therefore sometimes referred to as *naked-eye planets*.

Mercury, Venus, Earth, and Mars are nearer to the Sun and are collectively known as *rocky terrestrial* (Earth-like) planets. The rest four planets; Jupiter, Saturn, Uranus, and Neptune are further away from the Sun and are collectively known as *Jovian* (Jupiter like) planets. Unlike the Rocky terrestrial planets Jovian planets are gas or ice giant planets. Figure 6.19 shows Terrestrial and Jovian planets. The differences between Rocky Terrestrial and Jovian planets are shown in Table 6.1.

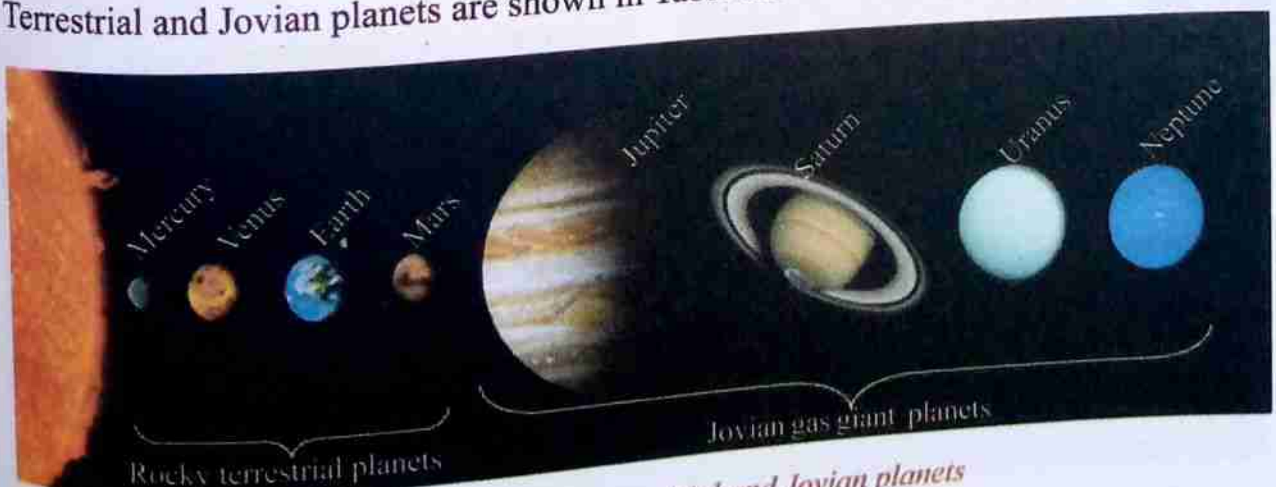


Figure 6.19: *Rocky terrestrial and Jovian planets*

Table 6.1: *Differences between Rocky terrestrial and Jovian planets*

Rocky terrestrial planets	Jovian planets
They are mostly made up of solid materials	They are mostly made up of gases or ice
They are smaller in size compared to Jovian planets	They are larger in size than the rocky terrestrial planets
They are closer to the Sun	They are far away from the Sun
They are denser than Jovian planets	They are less dense
They have fewer moons	They have many moons

Among the near planets, Venus reflecting light from the Sun, shines brightly like a star in the sky for some hours before it disappears. The planet disappears because it lies between the Sun and the Earth. Venus is nicknamed as *morning star* and *evening star* (Figure 6.20). When Venus is east of the Sun it sets after the Sun sets and it is known as the evening star. When it is west of the Sun, it rises before the Sun and is known as morning star.

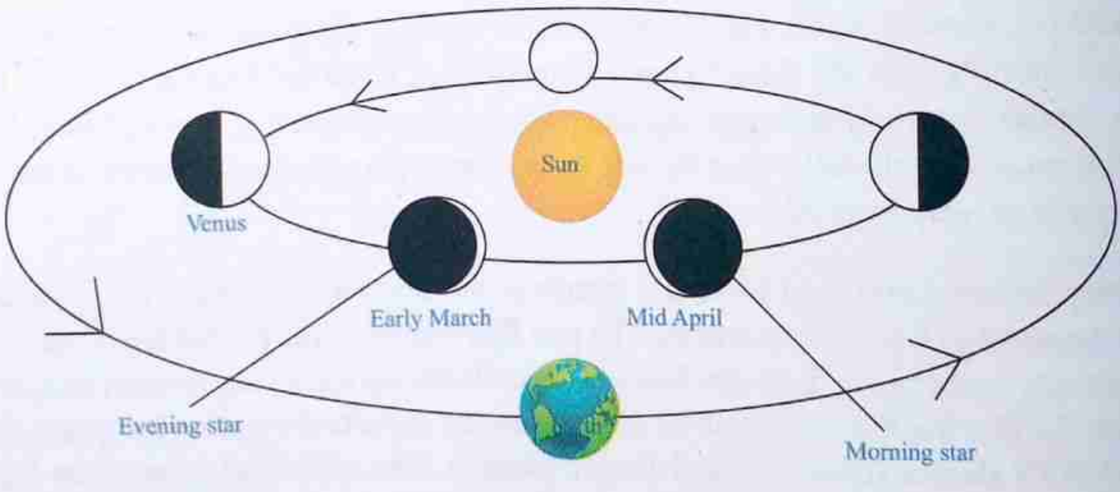


Figure 6.20: *Position of Venus with respect to Earth and the Sun*

Task 6.2

(a) Using internet and library search as a guide, write brief notes on how you can observe the naked-eye planets at your locality.

(b) Plan and observe the planets during several clear nights.

(c) Discuss the number of planets you were able to view during each night.

(d) Which day had more visible planets? Why?

Interior structure of the planets

The interior regions of planets are composed of different materials. The terrestrial planets consist mainly of rocky and metallic materials. Since they were once in molten state, their interior structures are arranged according to the densities of the materials. On the other hand the interior composition of Jovian planets are mainly gases and other elements as depicted in Figure 6.21.

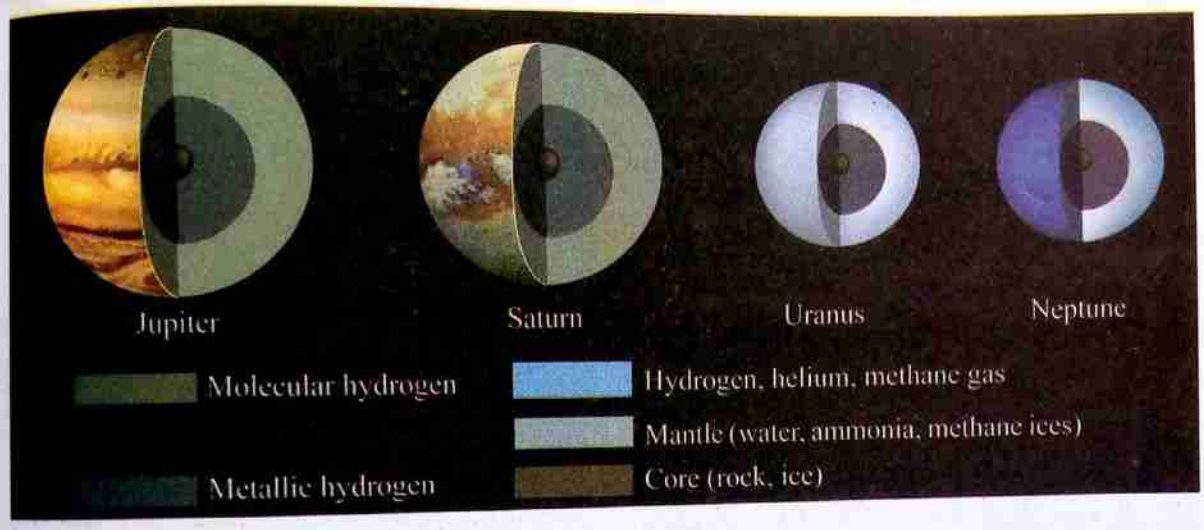


Figure 6.21: *Interior models of Jovian planets, showing some features and chemical compositions*

The Earth is currently the only planet known to be habitable for human beings. However, several studies are underway to explore the possibilities of other habitable planets including planets that are outside the solar system.

Task 6.3

1. In groups, discuss the possible elements found on the Earth.
2. Present the results in the class.

Other planets

There are other planets that revolve around stars beyond our solar system; such planets are known as exoplanets. For example, WASP-71b is a Tanzanite-exoplanet that revolves around a Mpingo-star named WASP-71 and was discovered in the year 2012. Figure 6.22 shows the exoplanet and star system.

Figure 6.22: *Tanzanite-exoplanet (WASP-71b) revolving around Mpingo-star (WASP-71).*

The Tanzanite exoplanet is more massive and larger than the Earth and even Jupiter. The comparison between the exoplanet, the Earth and Jupiter is shown in Figure 6.23.

Source: *https://wasp-planets.net/naming/wasp-71/*

Figure 6.23: *The Tanzanite-exoplanet, WASP-71b compared to Jupiter and the Earth*

Characteristics used to name a celestial object as a planet are:

(a) It should orbit a star.

(b) It should be massive enough so that its own gravity causes it to assume a spherical shape.

(c) It should clear the neighborhood around its orbit.

Pluto fails to meet the third characteristic, that is, it has not cleared neighbourhood around it. Therefore, it is termed as a dwarf planet. Table 6.2 shows the differences between stars and planets.

Table 6.2: *Differences between stars and planets*

Stars	Planets
Emit their own light	Do not emit their own light
Twinkle at night	Do not twinkle at night
Appear to be moving from east to west	Planets in our solar system move around the Sun from west to east
Their surface temperatures are usually very high	Their surface temperatures depend on their distances from the star that they orbit and the size of the atmosphere they have
In the universe, stars are countless in number. In our solar system, there is one star, the Sun	There is a countable number of planets, only eight in our solar system
Very big in size as compared to planets but appear small because they are very far from the Earth	Very small in size as compared to stars

Task 6.4

Using any relevant reference materials, draw a table that indicates the following parameters: distance (in km) of planets in the solar system from the Sun, number of natural satellites for each planet, the estimated density, temperature and period to orbit the Sun, the availability of atmosphere, colour and availability of water.

Other celestial objects in the solar system

Asteroids are small bodies which move around the Sun. They are sometimes called minor planets found between Mars and Jupiter along an asteroid belt (Figure 6.24). Asteroids are smaller than the planets and can only be seen through large telescopes.

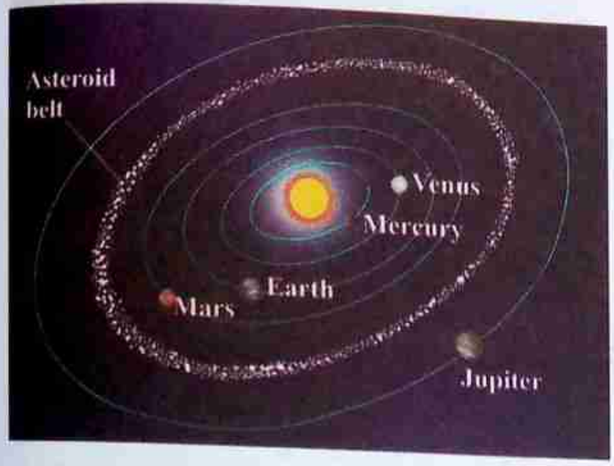


Figure 6.24: *Asteroid belt*

Meteors and meteorites

At night, when the sky is clear and the Moon is not shining, you may sometimes see bright streaks of light crossing the sky. The streaks of light are caused by space rocks that enter the Earth's atmosphere from outer space. Most of them are fragments of asteroids that broke apart long time ago in the asteroid belt. Once in the Earth's atmosphere, these rocks are called *meteors* (Figure 6.25). Although meteors are not stars, they are commonly known as *shooting stars*. This is because, as meteors cross the Earth's atmosphere at very high speed they burn up due to the heat generated by friction with the atmosphere. A meteor therefore glows and evaporates quickly. However, some meteors are large and so they can reach the Earth before they evaporate completely. The one that reaches the Earth is called a *meteorite*. A meteorite found at Ndolezi-Mbozi, in Songwe Region, Tanzania is a good example of a meteorite.

Figure 6.25: *Meteor moving towards the Earth*

Comets

Comets are also members of the solar system. They revolve around the Sun in highly elliptical orbits. However, their period of revolution around the Sun is usually very long. Comets appear generally as objects with bright heads and long tails. The tail of the comet is always directed away from the Sun (see Figure 6.26). Most of the comets were formed from condensed interstellar gases and dust in the early stages of the formation of the universe.

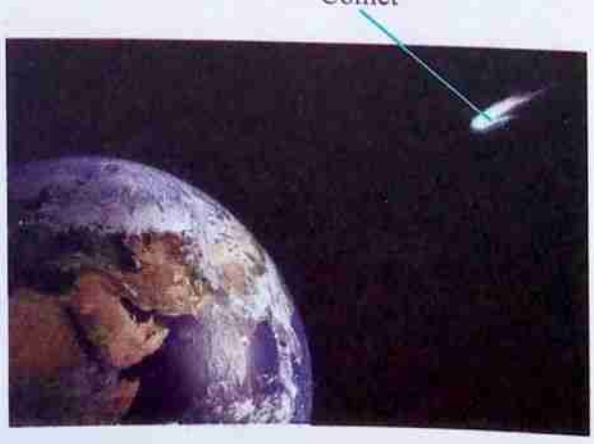


Figure 6.26: *Comet moving towards the Earth*

Zodiacal light

After sunset and before sunrise, scattering of sunlight caused by interplanetary dust particles can be observed. The scattered light may appear as a visible cone of faint, diffuse, and roughly triangular white glow in the sky called Zodiacal light as illustrated in Figure 6.27.

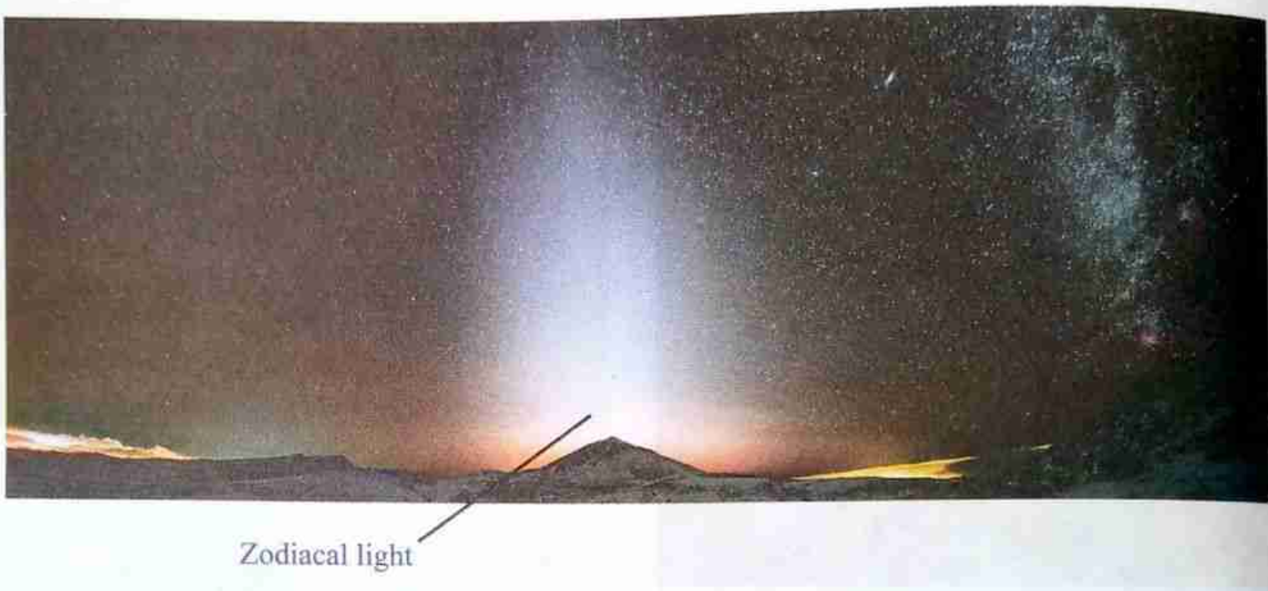


Figure 6.27: *Zodiacal light*

Activity 6.3

Aim: To observe the night sky.

Materials: Night sky, sheets of white paper, pencil, compass, eraser

Procedure

1. Observe the sky at least twice in one evening, with an interval of about two hours between observations.
2. Watch the Moon and note its position relative to the stars. Then write your observations.
3. After one or two hours, look again and note the new position of the Moon relative to the stars.
4. Extend this experiment to a month. Note the position of the Moon at the same hour on each possible night for a month.

Note: The observations should relate to the stars, naked-eye planets, constellations, Moon, galaxies, zodiacal light, and also to the position in the sky relative to the horizon.

Question

(a) What did you learn from your night-sky observations?

(b) Why do stars twinkle at the night sky?

(c) Why are stars not visible during the day?

Exercise 6.3

1. Briefly explain two theories used to describe the structure of the solar system. What is their significance?

2. The Sun as the nearest star to the Earth forms a solar system.
   (a) Name the celestial bodies which make up the solar system.
   (b) What force prevents the solar system from breaking apart?
3. (a) What are the differences between rocky terrestrial and Jovian planets?
   (b) Where are the Jovian planets found in the solar system?
4. Distinguish between morning star and evening star.
5. (a) What is meant by an exoplanet?
   (b) Draw (not to scale) the Mpingo-star and its Tanzanite-exoplanet system.
6. Why is Pluto no longer regarded as a planet?

Gravitational force

You may have asked yourself questions such as: why the Moon, planets, and other distant objects don't leave their orbits and fall down? What helps the Sun to hold all the planets in the solar system? Such questions were also asked by our ancestors. These are fundamental questions whose answers have led to better understanding of the solar system and the universe at large. Generally, celestial objects are kept in their orbits without flying off in a straight line by gravitational force.

In the solar system all the planets go around the Sun in elliptical paths. This indicates that, there exists a force between the Sun and the planets. Sir Isaac Newton recognized that, the force that attracts objects towards the Earth is the same for the Moon, the planets, and the other celestial bodies. That force of attraction between objects is called the gravitational force. Newton formulated a law of universal gravitation which states that:

Every object in the universe attracts every other object with a force which is directly proportional to the product of their masses and inversely proportional to the square of the distance between their centres.

The gravitational force is along the line joining the centres of the objects as illustrated in Figure 6.28.

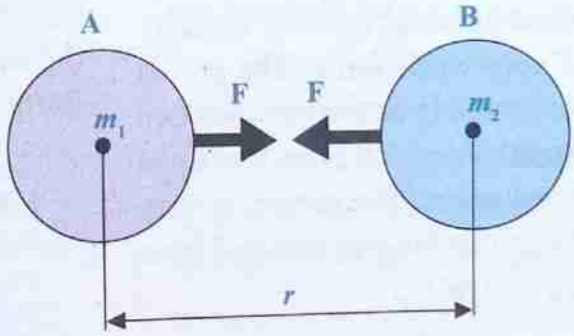


Figure 6.28: *Gravitational force between two uniform objects*

Let two objects A and B of m_1 and m_2 lie at a distance r between their centres as shown in Figure 6.28. Let the force of attraction between the two objects be F. According to the universal law of gravitation, the force between the two objects is directly proportional to the product of their masses. That is,

$$F \propto m_1 m_2,$$

and the force between the two objects is inversely proportional to the square of the distance between their centres, that is,

$$F \propto \frac{1}{r^2}$$

Combination of the two equations yields:

$$F \propto \frac{m_1 m_2}{r^2}$$

Therefore,

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

where G is a constant of proportionality called the *universal gravitational constant*.

Gravitational force is actually a very weak force. The pull is too weak to be felt between two small objects. It is only when at least one of the masses is very large that the gravitational force can be felt.

Acceleration due to gravity for an object on the Earth's surface

When an object is placed within the Earth's gravity, it experiences a gravitational force towards the centre of the Earth. This force causes the object to accelerate towards the Earth's centre with an acceleration known as acceleration due to gravity, g. Since the radius of the Earth increases slightly from the poles to the equator, (because of the bulge at the equator) the value of g for an object on the Earth's surface becomes slightly greater at the poles than at the equator. For most calculations, g is taken to be more or less constant on or near the Earth's surface. The acceleration due to gravity for an object on the Earth's surface is given as,

$$g = \frac{GM_E}{R_E^2}$$

where G is the gravitational constant, M_E is the mass of the Earth and R_E is the radius of the Earth.

Acceleration due to gravity varies from one planet to another as well as from one satellite to another. The gravity on the Moon is 1/6 times the gravity on the Earth. As a result, an object on the Moon feels less weight compared to its weight on the Earth. Since the mass of an object is constant everywhere, its weight changes depending on the value of gravity at a given place in the universe. Figure 6.29 illustrates the variation of weights for a given person while on Earth and assumed to be on Mars and Moon.

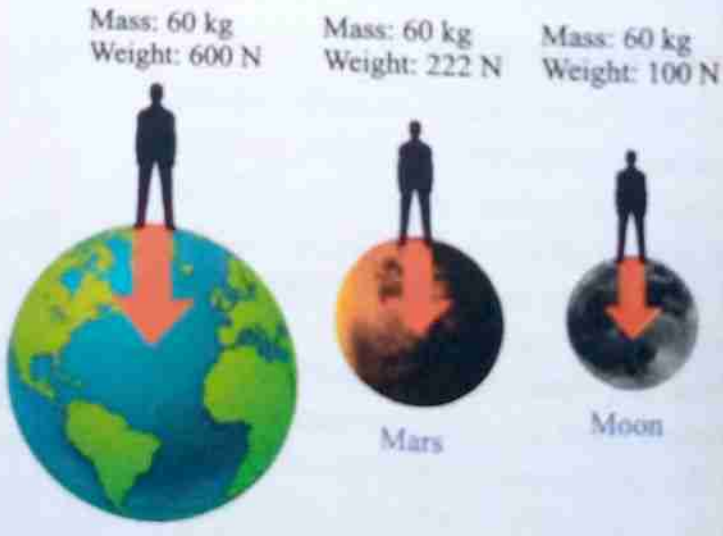


Figure 6.29: *Weight of a person varying with gravity*

Task 6.4

Using the values given in Figure 6.29, estimate the value of acceleration due to gravity on the Earth, Mars and Moon.

Exercise 6.4

1. What is the difference between gravity and gravitational force?
2. State Newton's universal law of gravitation and list the quantities and the units involved.
3. Describe why the weight of an object on the Earth is larger than the weight of the same object placed on the Moon and Mars.

The Earth and its Moon

The Earth and its Moon form a unique pair in the solar system. The Moon of the Earth is the fifth largest moon in the solar system. It has a diameter of about 3 476 km and a mass of about 7.35×10^{22} kg. As it has been explained in the previous section, besides the Earth, the Moon is another body in the solar system upon which humankind has landed (see Figure 6.30). The Moon is a natural satellite that moves in an elliptical orbit around the Earth. The Moon is held in a fixed orbit by the gravitational force between the Earth and itself.

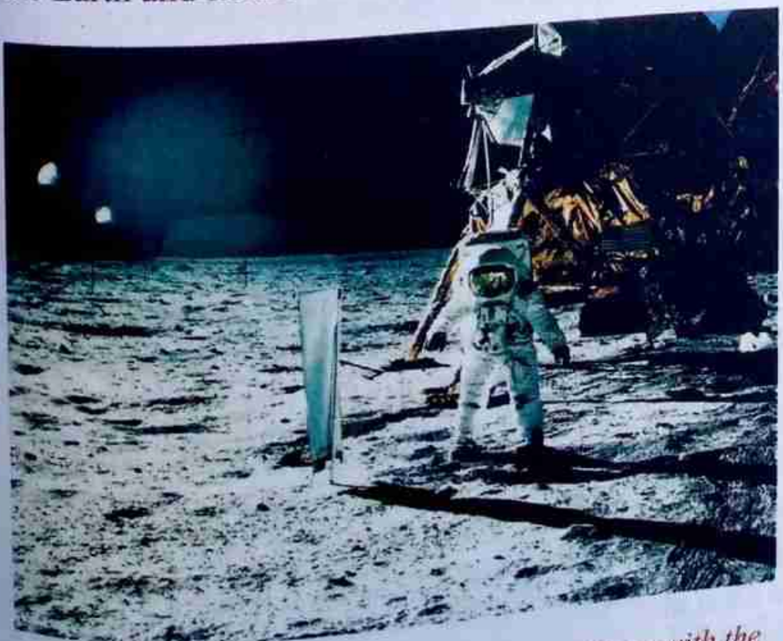
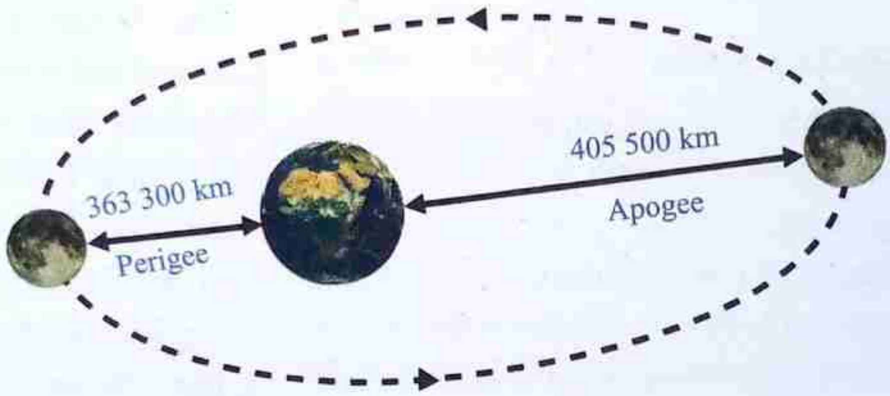


Figure 6.30: *Astronaut Buzz Aldrin on the Moon with the Apollo 11 spacecraft in the year 1969*

Like the Earth, the Moon has an iron core surrounded by a rocky mantle and crust. Unlike the Earth, no part of the Moon's iron core is molten so it does not have a magnetic field. Surface gravity on the Moon is 1/6 times that of the Earth.

The Moon revolves in an anticlockwise direction around the Earth in an elliptical orbit. The Moon's orbit is tilted at 5° relative to the Earth's orbit around the Sun (see Figure 6.31 (a)). The distance between the Earth and the Moon varies depending on the position of the Moon as it revolves around the Earth. The nearest distance of the Moon from the Earth is known as *Perigee*, and it is about 363 300 km. On the other hand, *apogee* refers to the distance between the Earth and the Moon when the Moon is furthest from the Earth, and it is about 405 500 km (Figure 6.31 (b)). The average distance is 384 000 km. It takes about 27.3 days for the Moon to complete one orbit. This period of time is called the *Sidereal month.*

(a)



(b)

Figure 6.31: *(a) Moon orbit's the Earth (b) perige and apogee distances*

As the Moon revolves around the Earth, it also rotates on its own axis but the time it takes to rotate once about its axis is equal to the time it takes to revolve around the Earth. Consequently, the same side of the moon faces towards the Earth throughout the year. The side facing the Earth is called near side while the side away from the Earth is called far side. The spinning of the Earth causes the Moon to rise and set each day, just like the Sun. However, because of the Moon's orbital motion around the Earth, each day the Moon rises about 50 minutes later than the previous day. As a result, the Moon can be seen at different times of the day and night within a month.

Surface features of the Moon

The Moon is not like the Earth as it does not have oceans, lakes, rivers, streams, wind-blown and ice field at its poles. Major features of the Moon's surface can be seen by just looking up at it. It has bright and dark areas. These distinctive terrains are the bright lunar highlands, also known as the lunar terrae. Terrae is the Latin word for 'land'. The areas have many craters and are covered with a high refractive layer of fine dust. The highlands are geographically the oldest parts of the Moon's surface. The dark plain is called the lunar maria, Latin for 'sea' which resembles the ideas of

Thomas Hariot and Galileo Galilei, who were the first scientist to examine the Moon with telescopes. The plains filled with dark solidified lava are less cratered than the highlands (Figure 6.32).

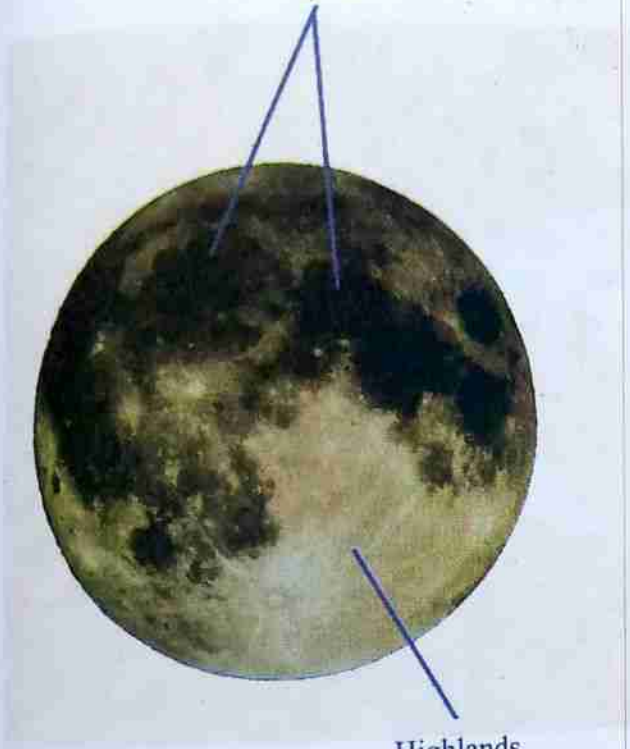


Figure 6.32: *Moon's highlands and maria*

The maria, which make about 16% of the Moon's surface, are huge impact craters that were later flooded with molten lava. Most of the maria is covered with regolith, a mixture of fine dust and rocky debris produced by meteor impact.

Task 6.5

Observe the Moons' cycle for a month, then using internet or library search:

(a) Draw the phases of the Moon throughout the monthly cycle.

(b) Explain the phases of the Moon throughout the monthly cycle.

Effects of gravitational force between Earth-Moon-Sun

Referring to Newton's law of universal gravitation, there exists a gravitational force of attraction between the Earth and the Moon. This force is more significant on the smaller mass, in this case the Moon. However, since the Moon pulls the Earth towards its centre, less dense materials such as fluid, particularly water bodies on the Earth tend to move in the direction of the Moon's gravitational pull. This causes ocean water rise on the side facing the Moon, and fall on the opposite side, resulting to ocean tides.

Ocean tides

Tides are periodic rise and fall of ocean water level caused by the combined effects of the gravitational forces exerted on the Earth by the Moon and the Sun, and the rotation of the Earth.

Generally, tides are periodic rise and fall of ocean level due to combined effects of gravitational interaction between the Earth-Moon-Sun. Since the distance between the Earth and the Moon is shorter compared to that between the Earth and the Sun, tides caused by the Moon are stronger than those caused by the Sun.

How tides occur

Moon's gravitational pull on Earth causes the oceans on the Earth's side facing the Moon to bulge out in the direction of the Moon to bulge out in the direction of the

Moon. This results to the oceans on the opposite side to bulge in. Therefore, ocean water levels fluctuate daily as the Sun, Moon and Earth interact. As the Moon travels around the Earth and Earth travel around the Sun, the combined gravitational forces cause the ocean water levels on Earth to rise and fall. Since the Earth is rotating while this is happening, two tides occur each day at a given location. There are *high* and *low tides* (Figure 6.33).

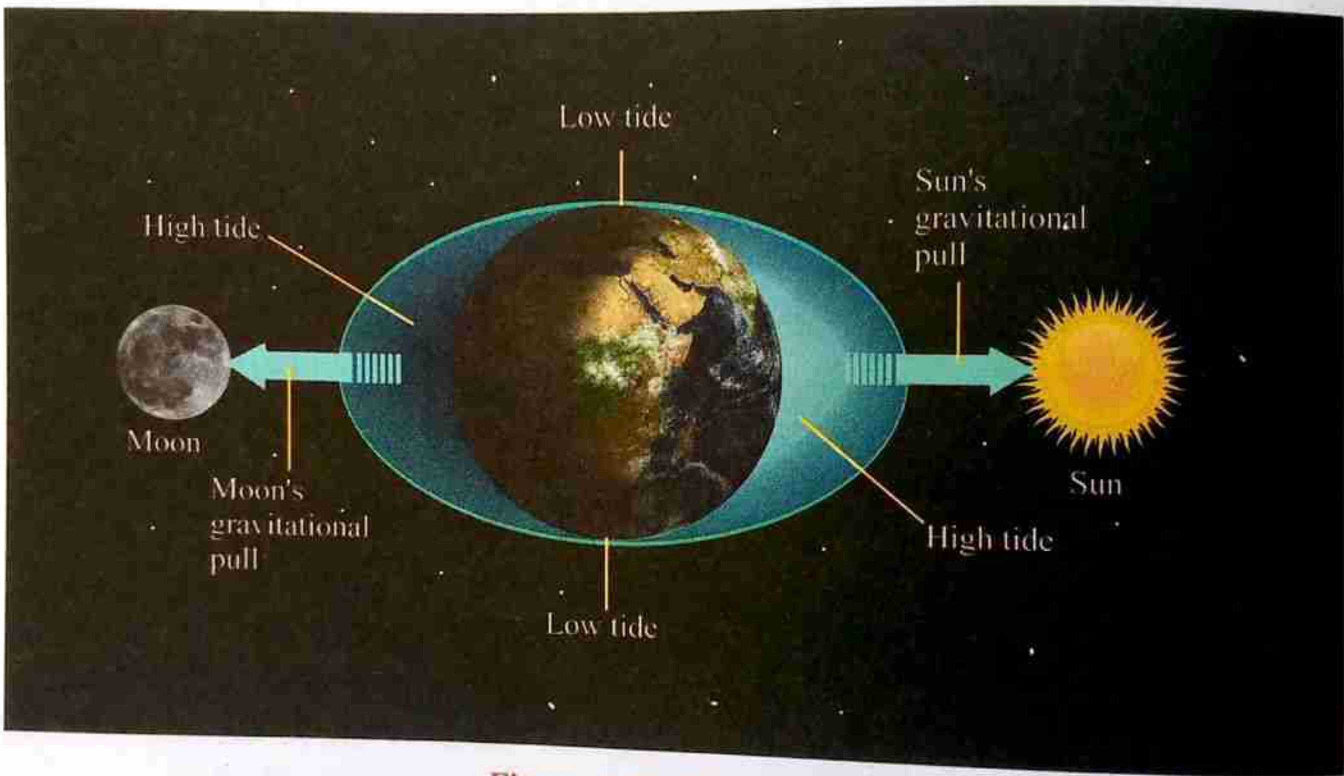


Figure. 6.33: *Ocean tides*

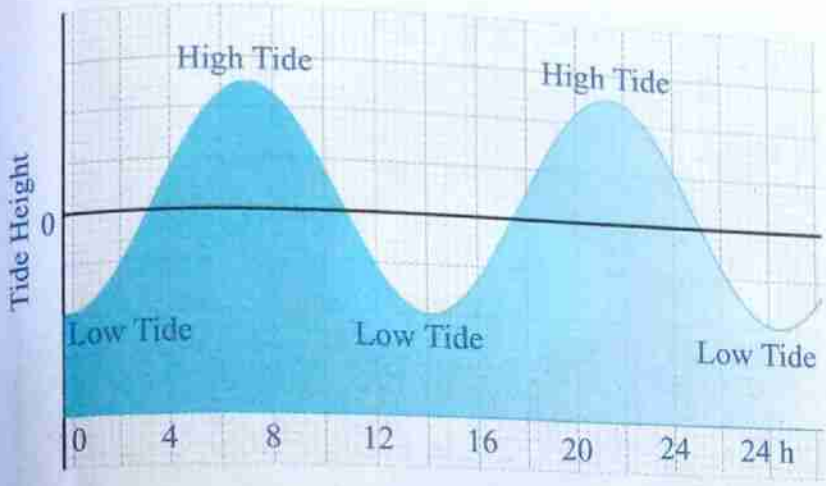
Types of tides

Tides can be grouped into various types based on their respective height and lunar phase.

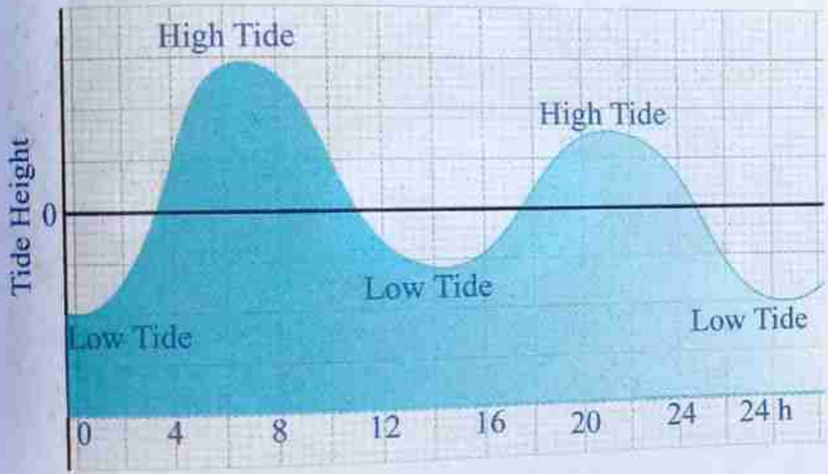
Classification based on height

Based on height, tides are classified as *high tide* and *low tide*. High tide occurs when ocean water reaches its highest height within the tide cycle, while low tide occurs when ocean water reaches its lowest height within the tide cycle.

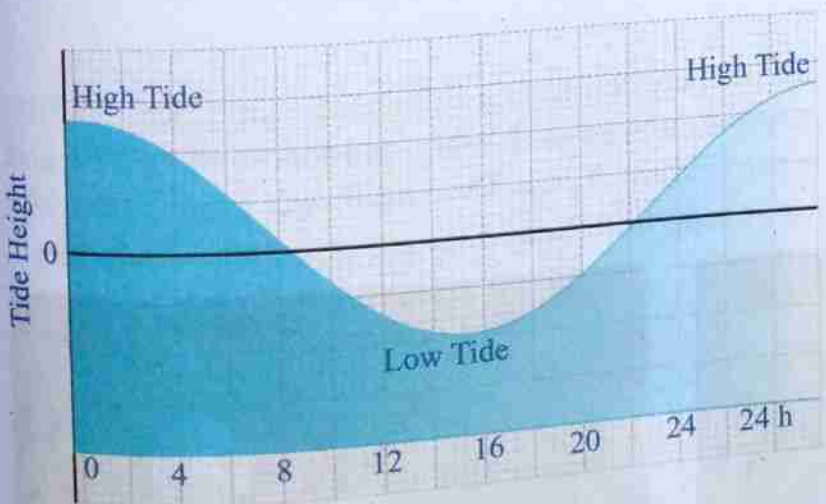
The ocean is constantly moving from high tide to low tide, and then back to high tide. Generally, most areas experience two high tides and two low tides for every lunar day. When there are two high tides and two low tides of about the same height, the pattern is called a *semidiurnal tide* (Figure 6.34 (a)). If the high and low tides differ in height, the pattern is called a *mixed semidiurnal tide* (Figure 6.34 (b)). However, some areas experience only one high and one low tide each day; this pattern is called *diurnal tide* (Figure 6.34 (c)).



(a) Semidiurnal tide



(b) Mixed semidiurnal tide



(c) Diurnal tide

Figure 6. 34: *Tidal cycles*

Note:

1. There is a time interval of about 12 hours and 25 minutes between the two high tides.
2. A day with 24 hours is called a solar day while a day with 24 hours and 50 minutes is called a lunar day.

Classification based on lunar phase

The height of rising oceanic water varies greatly depending upon the position of the Sun and the Moon with respect to the Earth. In the lunar phase, tides are classified as *spring tides* and *neap tides*.

Spring tides

These tides occur during the new moon and full moon phases when the Sun, the Moon and the Earth are aligned. Their combined effect produces strong spring tides, and height of the tide will be higher. Spring tides occur twice a month. That is during full moon period and new moon period as illustrated in Figure 6.35.

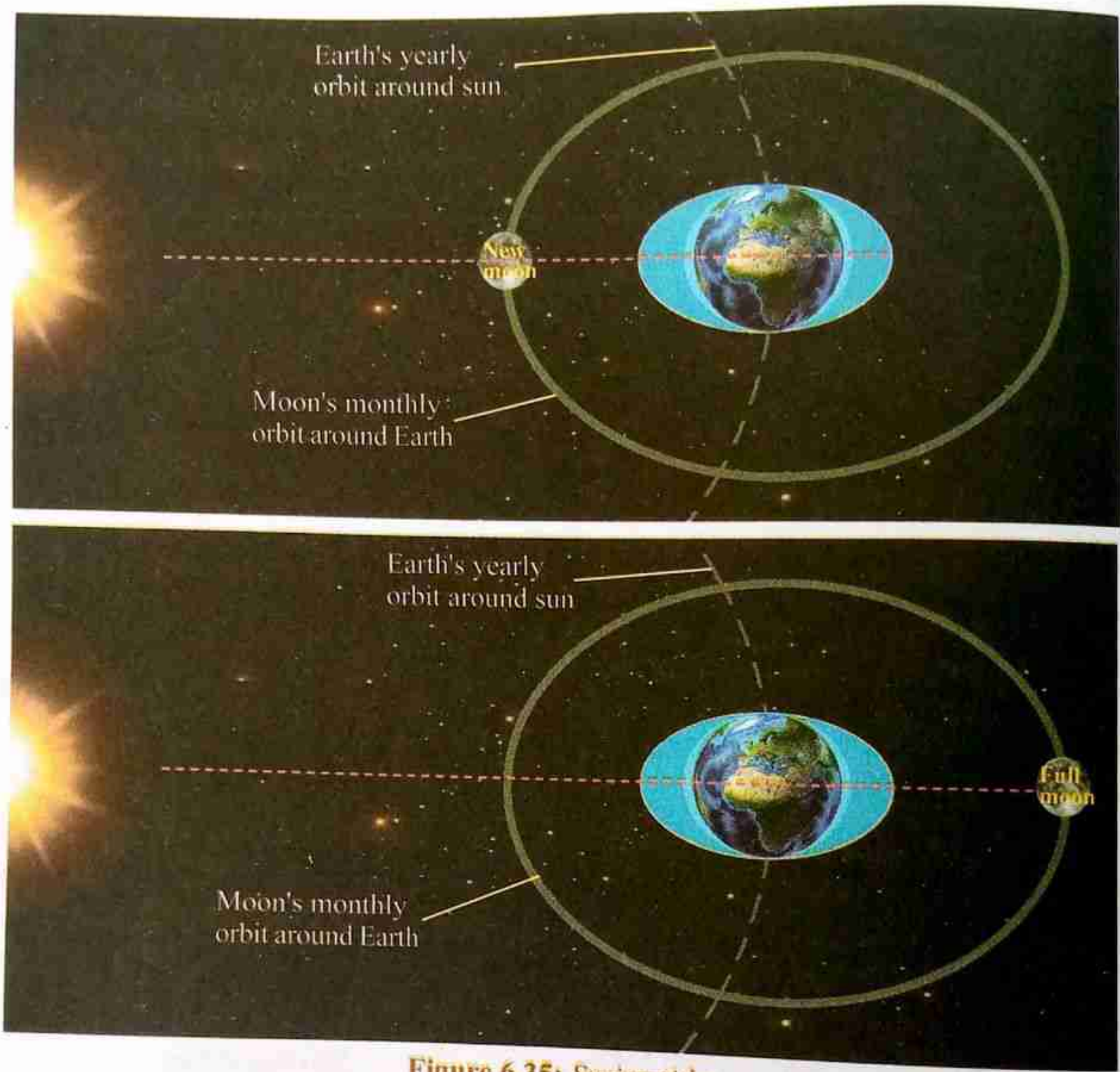
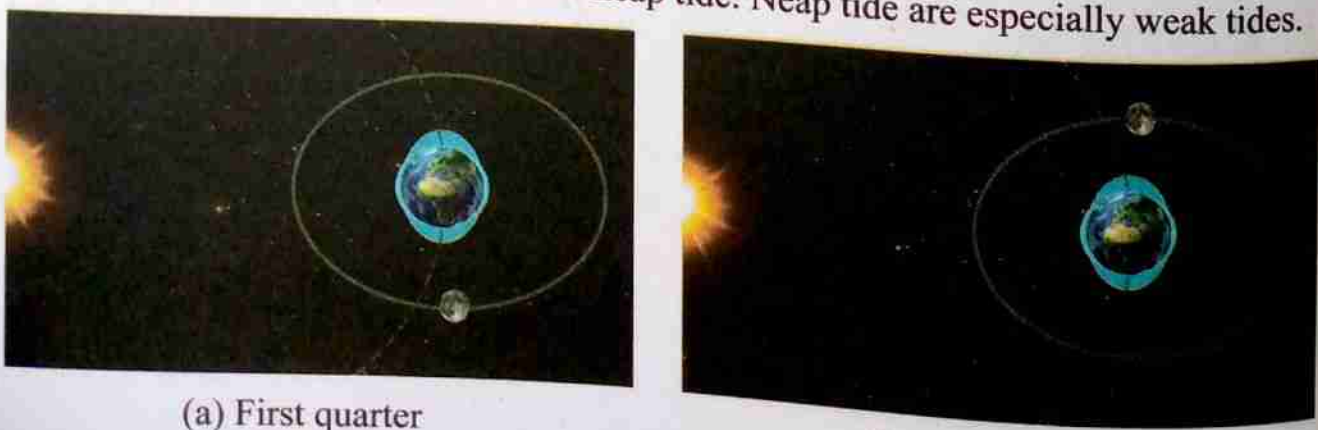


Figure 6.35: *Spring tides*

Neap tides

Normally, there is a seven-day interval between spring tides and neap tides. During this time, the Moon and Sun are perpendicular to each other and forces of the Sun and Moon tend to counteract one another (Figure 6.36). The result is a small difference between high and low tides, known as neap tide. Neap tide are especially weak tides.



(a) First quarter

(b) Third quarter

Figure 6.36: *Neap tides*

Importance of tides

1. High tides help in navigation. They raise the water level close to the shores. This helps ships to move to harbours more easily.
2. High tides also help in fishing. More fish come closer to the shore during the high tide. This enables fishermen to get plentiful catch.
3. Tides are also used for the generation of tidal electricity.

Exercise 6.4

1. With the aid of a diagram describe the following terms as used in astronomy:
   (a) Apogee
   (b) Perigee
   (c) Lunar highlands
   (d) Maria
2. Explain the effects of gravitational force between Sun and Earth or Sun and Moon.
3. Distinguish between the following:
   (a) High tides and low tides
   (b) Neap tides and spring tides
   (c) Semidiurnal tides and mixed semidiurnal tides
4. Explain why high tides occur twice a day?
5. With reasons, explain when the spring tides and neap tides are experienced through the Moon cycle.

Chapter summary

1. Astronomy is a science that deals with the study of objects and phenomena beyond the Earth's atmosphere.
2. In early times, astronomy was used to measure time, predict seasons and assist in navigation.
3. Our solar system is made up of eight planets, the Sun, thousands of asteroids, countless comets and meteoroids.
4. Stars are giant spheres of hot gases called plasma.
5. A galaxy is a large group of stars, dust and gas held together by mutual gravitational force.
6. Gravitational force is the attractive force that exists between any two objects that have mass.
7. A constellation is a group of stars that form a definite pattern in the sky when viewed from the Earth.
8. Constellations are used in religion, agriculture and in navigation.
9. The Earth has one Moon. The Moon of the Earth is the fifth largest moon in the solar system.
10. Tides are periodic rise and fall of ocean due to combined effect of gravitational interaction between the Earth, Moon and Sun.
11. Based on height, tides are classified as high tide and low tide.

12. High tide occurs when ocean water reaches its highest height within the tide cycle, while low tide occurs when ocean water reaches its lowest height within the tide cycle.
13. Based on lunar phase, tides are classified as spring tides and neap tides.
14. Spring tides occur during the new moon and full moon phases, when the Sun, Moon and Earth are aligned.
15. Neap tides occur during the first and third quarter of the moon when the Sun and the Moon are perpendicular to each other with respect to the position of the Earth.

Revision exercise 6

(Assume the value of $g = 10$ N/kg)

1. Choose the correct answer for items (a) and (b).
   (a) Which of the statements (i) to (iv) differentiate a planet from a star?
      (i) A planet is a celestial body and a star is a luminous celestial body.
      (ii) A star is a celestial body and a planet is a luminous celestial body.
      (iii) A planet has larger mass than a star.
      (iv) A star orbits a planet.
   (b) What is the Milky Way?
      (i) A planet in the solar system.
      (ii) A galaxy in which the solar system belongs.
      (iii) One of the solar systems.
      (iv) One of the stars in the solar system.
2. How are the bodies in the solar system kept in their normal position?
3. (a) What is a constellation?
   (b) List down any three constellations.
4. (a) What is a tide?
   (b) Briefly explain how tides are formed.
5. (a) What are the five planets that can be seen with naked eyes?
   (b) Distinguish between rocky terrestrial and Jovian planets.
   (c) Describe the composition, size and location of terrestrial planets and Jovian planets in the solar system.
6. Briefly explain how astronomy gave rise to the twelve months of the year.
7. How long does light take to travel from:
   (a) the Sun to Earth?
   (b) the Moon to Earth?
   (c) the centre of the Milky Way galaxy to the Earth?

8. Assume a body has mass of 40 kg. How much mass and weight will it have on the Moon?

9. A piece of geological mineral whose dimensions are 2 m × 4 m × 8 m found on the Moon weighs 36 N. Determine its weight and density when brought on the Earth.

10. Briefly explain the following terms:
   (a) Light pollution
   (b) Astro-tourism
   (c) Meteoroid
   (d) Meteorite
   (e) Shooting star

11. Distinguish between the following:
   (a) Galaxy and the universe.
   (b) Astronomical unit and light-year.

12. With the aid of a diagram distinguish between the celestial objects named WASAP-71b and WASAP-71.

13. With the aid of sketches, distinguish between the Gemini and the Orion constellations.

14. With examples, explain any three criteria that can be used to identify a constellation.

15. Distinguish between chromosphere and photosphere.

16. Which stellar (star) parts are the sources of solar winds, gamma radiations, flares and prominences?

17. Assume the Sun suddenly disappears. State any three demerits to humankind based on this disappearance.

18. Write any two electromagnetic radiation types that can be observed by a space-based telescope and any two that can be observed by a ground-based telescope.

19. In which region of the solar system are comets and asteroids found?

20. With the aid of a diagram, explain when Venus is termed as a morning star and evening star.

21. Using Figure 6.37, answer the following questions.
   (a) What is the distance in light years between Venus and Earth?
   (b) How long does light take to travel from the Sun to Venus?
   (1 light year $= 9.46 \times 10^{12}$ km).

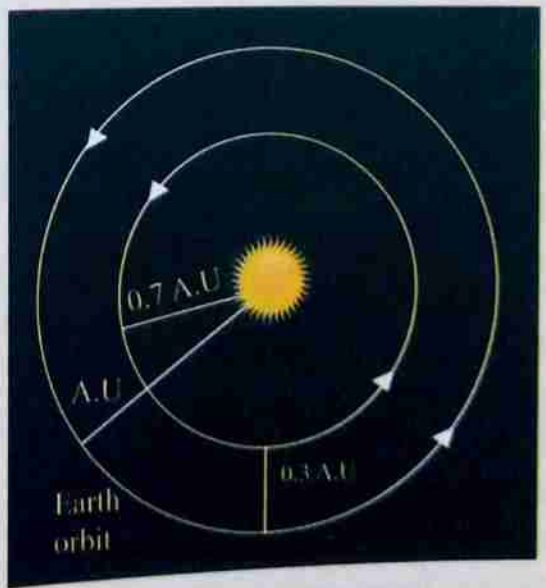


Figure 6.37

Chapter Seven

Physics of the Earth and its atmosphere

Introduction

In your life you may have experienced geohazards such as earthquakes, volcanic eruptions, tsunamis and landslides. These geohazards may have negative impacts to humans, other living organisms and environment. It is therefore necessary to understand the sources, signs and risks associated with these geohazards so as to be able to take precautions. Moreover, exploring for geoenergy, water, oil, gas and minerals needs the knowledge of processes taking place in the interior of the Earth and other atmospheric phenomena. Besides, the knowledge of the Physics of the Earth's atmosphere is important for maintaining temperature of the Earth. In this chapter, you will learn about the concept of geophysics, the structure and composition of the Earth, earthquakes and volcanoes, structure and composition of the atmosphere as well as the greenhouse effect and global warming. The competencies developed will enable you to to characterise structure of the Earth's interior and atmosphere, identify geohazards and risks posed by them. You will also be able to apply geophysical knowledge to manage your local environment.

The concept of the Earth and its atmosphere

The planet Earth in which humans live is subdivided into four spheres. These spheres are geosphere, hydrosphere, biosphere and atmosphere. The geosphere is part of the Earth which is solid (containing rocks and minerals). The solid Earth is surrounded by a gaseous layer called atmosphere. The atmosphere is held to the surface of the Earth by gravity. The hydrosphere is part of the Earth containing liquid, vapour and ice. The last component is the biosphere which contains all living organisms. The focus of this chapter is on understanding the physics of the Earth (geophysics) and its atmosphere.

Geophysics

Have you ever asked yourself about how natural calamities such as volcanic eruptions and earthquakes occur? Understanding the occurrence of these calamities requires the knowledge of characterising the interior of the Earth using physics techniques applied at the surface of the Earth. Geophysics is the subject designed to provide knowledge to human beings about the mechanism of various geological phenomena occurring on and in the Earth and their impacts. The word Geophysics has two parts; Geo- means Earth and Physics- means scientific knowledge of matter, energy and their interactions.

Generally, geophysics deals with the application of principles, laws and methods of physics to study the Earth.

The structure and composition of the Earth

It is believed that, in the beginning, the Earth existed in a molten state, in spherical shape. Materials in the molten state were arranged according to their respective densities. Less dense materials were at the top (outer) while denser materials were sedimented at the bottom (inner). As the universe expanded, the temperature decreased and resulted to solidification of the Earth. The process of solidification started from the outer surface towards the inner parts. This resulted to the layered structure of the Earth. Therefore, the interior of the solid Earth can be described in terms of chemical composition, physical and mechanical properties. In terms of chemical composition and physical properties the Earth is composed of three main layers arranged in a concentric manner. These are the crust, the mantle and the core. Figure 7.1 shows the interior structure of the Earth.

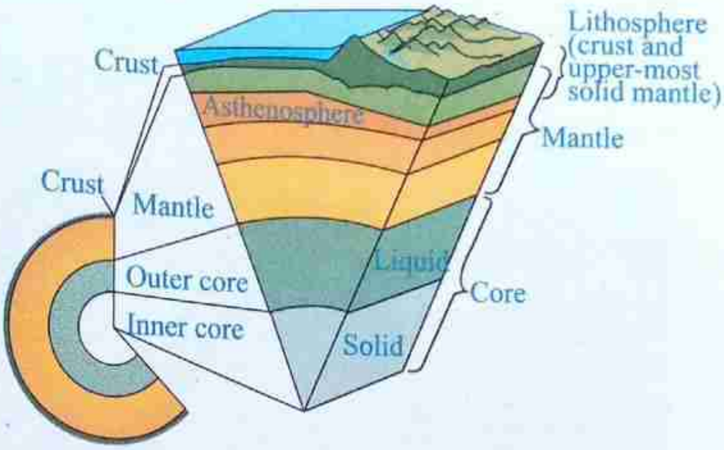


Figure 7.1: *Interior structure of the Earth*

The crust

The crust is the outer solid layer of the Earth. It is extremely thin, ranging from about 5 -100 km, as compared to the radius of the Earth which is about 6 371 km. The crust makes up less than 1% of the Earth by mass. It consists of continental crust and oceanic crust.

Continental crust is much thicker than oceanic crust. Under stable continental areas, the crust is about 35-40 km thick, and under mountain ranges it is often about 50-70 km thick. Under large mountain such as Mount Kilimanjaro the base of the crust can be as deep as 100 km. This layer is made up of rocks rich in silicate and magnesium minerals. The average density of the continental crust is about 2.7 $g\,cm^{-3}$.

Oceanic crust is the thin layer ranging from about 5-10 km thick, made up of rocks rich in iron and magnesium with low content of silica. Its average density is about 3.0 $g\,cm^{-3}$. The density of oceanic crust is somehow higher than that of the continental crust because of high contents of magnesium and iron than granite. Due to the differences in density, continental crust extends higher than the oceanic crust, which subducts underneath the continental crust. Figure 7.2 shows the continental and the ocean crust at converging point.

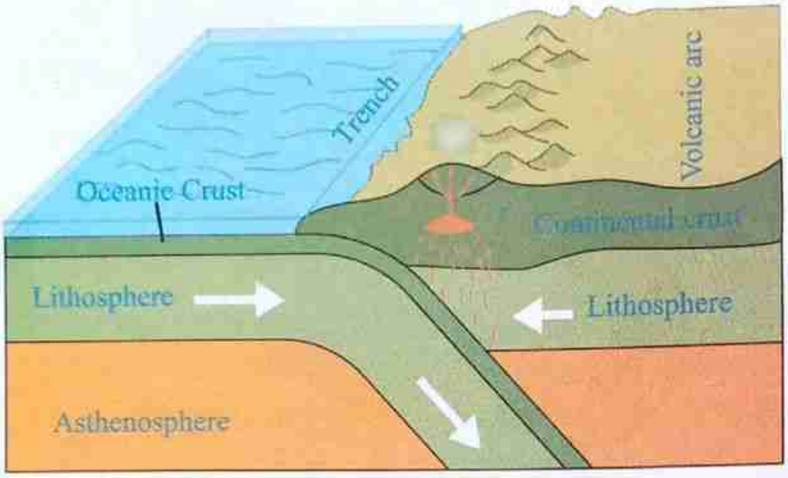


Figure 7.2: *Continental and the ocean crust*

The boundary between the crust and the mantle is called *Mohorovicic discontinuity* or simply Moho. The Moho zone ranges between one and several kilometres of thickness.

The mantle

Mantle is the thick layer between the Earth's crust and outer core. It begins from the Moho and extends to a depth of approximately 2 800 km below the Earth's surface.

The mantle makes about 82% of the volume and about 70% of the mass of the Earth, making it the Earth's thickest layer. It is composed of silicate rocks which are rich in iron and magnesium. The mantle is divided into two parts, the upper and lower mantle. The upper mantle has a temperature of about 1 000 °C. The temperature increases towards the centre to about 3 871°C near the core. Mantle rocks near the core are soft and able to move plastically. Thus, despite being solid rocks, the highly hot silicate material in the mantle near the core can flow over very long time frame (millions of years). Convection of materials in the mantle is the main mechanism of heat transfer from the core of the Earth to the outer regions of the Earth. It is the main force behind the continental movements as well as volcanism and earthquakes. The mantle-core boundary is called the *Gutenberg discontinuity* and is located at the depth of about 2 800 km.

The core

The core is the innermost part of the Earth. It extends from the Gutenberg discontinuity to the Earth's geometric centre. The core consists of two distinct regions, namely the outer core and the inner core. The outer core is composed of extremely hot iron and nickel. It extends from the lower mantle to a depth of about 5 100 km below the Earth's surface. Its temperature range is about 4 000 to 5 500 °C and it is about 2 300 km thick. Because of high temperatures, metals in the outer core are in molten form.

The inner core is composed of iron-nickel alloys. It is very small, making less than one percent of the volume of the Earth. It is solid because of the high gravitational pressure at this depth. It is separated from the mantle by the molten outer core, and is therefore, free to move independently.

On the other hand, based on mechanical (thermal) processes, the Earth can be divided into lithosphere, asthenosphere, mesosphere or lower mantle, outer core, and inner core as illustrated in Figure 7.3. These layers are based on how each layer responds to stress.

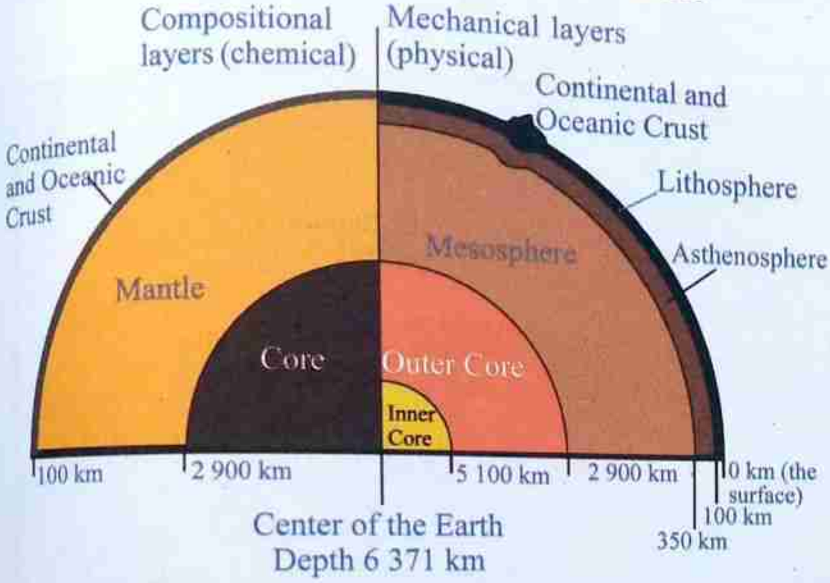


Figure 7.3: *Earth's interior based on mechanical process*

The lithosphere is the rigid outer layer of Earth, with "litho" meaning "rock." It consists of the crust and the solid outermost layer of the upper mantle. The lithosphere has both an oceanic and a continental component each topped by its own kind of crust. The oceanic lithosphere is thin and hard, with a thickness varying from 0 to 140 km. The continental lithosphere is more plastic and thicker. The lithosphere has a greater strength and lower density than the underlaying asthenosphere which enables tectonic plates to move.

The asthenosphere is the layer beneath the lithosphere that has characteristics of plastic, with 'astheno' meaning weak. The asthenosphere's most distinguishing feature is its ability to move. Because it is mechanically weak, it will flow and migrate over geologic time scales while still solid. Movement of this layer allows the lithospheric plate to move, which is partly caused by convection of strong internal heat.

The mesosphere, also known as the lower mantle, is more rigid and immobile relative to the asthenosphere, though it is still quite hot. This is due to the temperature increase with depth.

Other layers described in terms of mechanical properties are the *outer core* which is the only liquid layer found within Earth and the *inner core* which is in solid form due to high gravitational pressure.

Layers of the Earth's interior provide information on how the planet evolved, the possible layers that make up other planets, the source of Earth's resources like different minerals, and much more.

Task 7.1

In groups, discuss why the outer core is in liquid form while the inner core is in solid form.

Tectonic plates

The Earth's crust and part of the mantle, together referred to as the lithosphere, is cracked into huge pieces called *tectonic plates*. These plates float on top of the semi-molten rocks and therefore they move at a very slow speed. Some of the tectonic plates are moving apart and some are moving towards each other. The movements of tectonic plates tell us that some continents are either moving apart or towards each other. This process is referred to as *continental drift* and has been progressing for hundreds of millions of years. Tectonic plate movements have split the continents into many plates including the North American, Caribbean, South American, Scotia, Antarctic, Eurasian, Arabian, African, Indian, Philippine, Australian, Pacific, Juan de Fuca, Cocos, and Nazca as shown in Figure 7.4

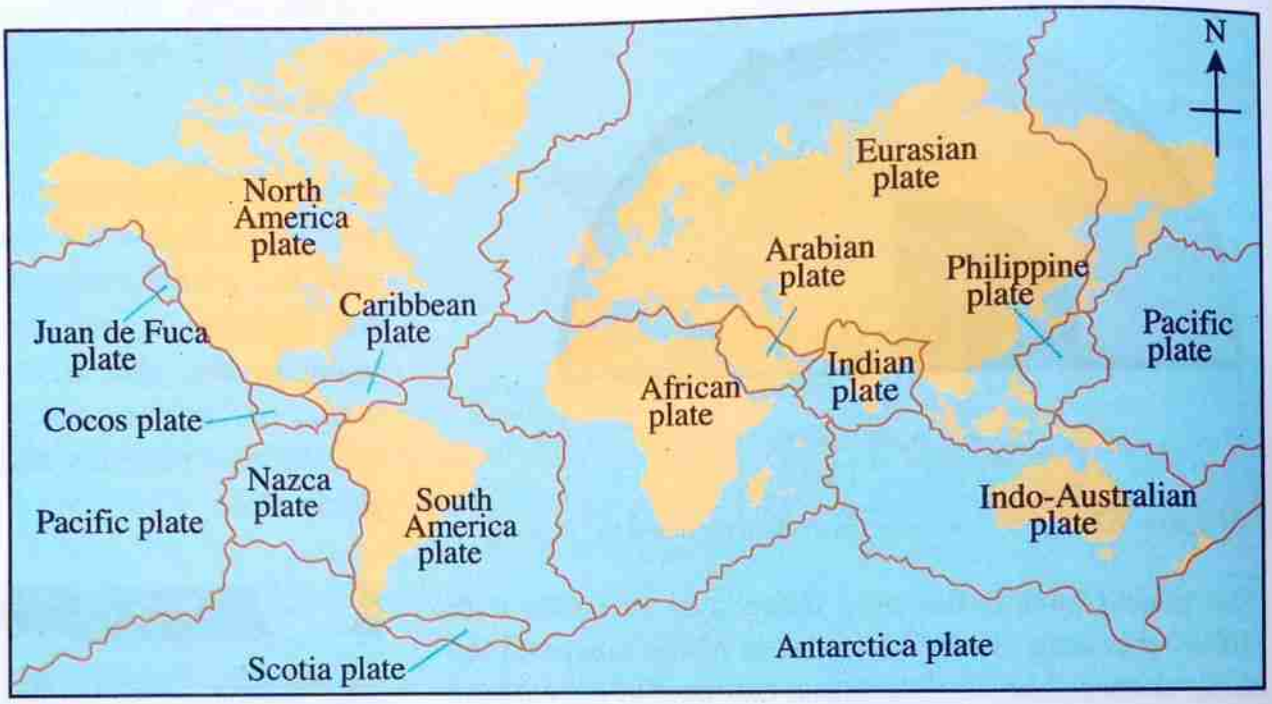
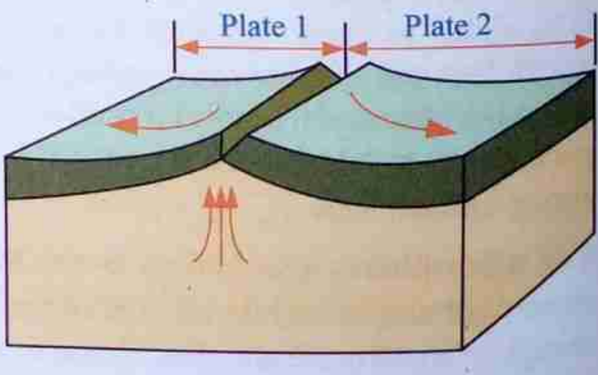


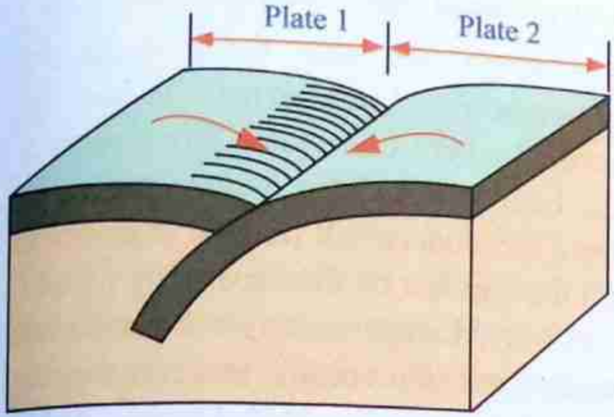
Figure 7.4: *Earth's tectonic plates*

The line where two tectonic plates meet is called a *boundary*. There are three main types of boundaries. These are *constructive boundaries, destructive boundaries*, and *conservative boundaries.*

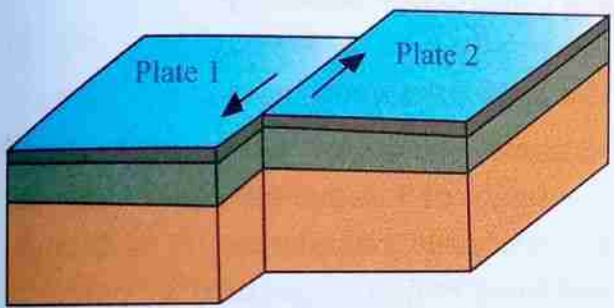
A constructive boundary, also called *divergent* or *accreting,* is formed at the edges of two plates moving away from each other as shown in Figure 7.5 (a). Destructive boundary, sometimes called *convergent plate margin* or *tensional plate margin,* occurs when two plates move towards each other as depicted in Figure 7.5 (b). Most of destructive boundaries are represented by a system of subduction zones in which one of the colliding plates descents into the mantle. Conservative boundaries are formed when two plates slide past each other in opposite directions as shown in Figure 7.5 (c), or in the same direction, but at different speeds. They are sometimes called *transform plate margins*.



(a) Constructive boundaries



(b) destructive boundaries



(c) Conservative boundaries

Figure 7.5: *Plate boundaries*

Exercise 7.1

1. Differentiate between oceanic crust and continental crust.
2. What is the main mechanism of heat transfer from the core of the Earth to the outer regions of the Earth?
3. Explain the compositions of different layers of the interior of the Earth.
4. Explain the importance of the three main layers of the solid Earth.
5. Describe the three main types of tectonic plate boundaries.
6. Among the three layers of the internal structure of the Earth, which is the densest layer? Why?

Earthquakes and volcanoes

Volcanoes and earthquakes are closely related. They are both caused by the movement of molten rock and heat deep inside the Earth. These movements are referred to as subterranean movements. Earthquakes and volcanic activities mostly happen near tectonic plate boundaries.

Volcanoes

Volcanoes are features in the Earth's crust that allow lava to escape from a large pool of molten rock beneath the Earth's surface. The molten rock is called *magma* and the pool of magma is called the *magma chamber.* Magma originates from the mantle, where high temperature and pressure cause the rock to melt. Magma in the magma chamber is less dense than the surrounding rock. Thus, the rock provides buoyant forces that tend to drive the magma upwards. If the magma finds a path to the Earth's surface, the result is a volcanic eruption. Magma that has been erupted to the Earth's surface is called *lava.* Figure 7.6 shows an example of volcanic eruption.

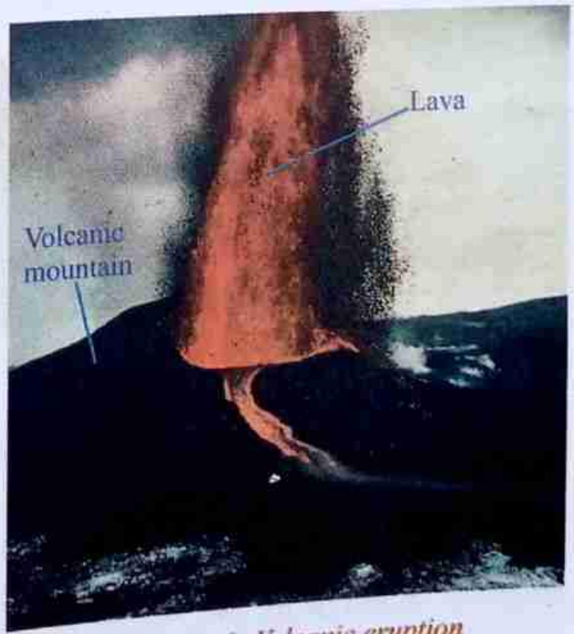


Figure 7.6: *Volcanic eruption*

Note that, a volcanic eruption refers to the release of lava, volcanic ash and/or gases from a volcano. Most volcanoes are situated over the magma chambers which are in many cases along the constructive and destructive boundaries of tectonic plates. However, some volcanoes may occur far from plate boundaries.

Volcanoes at destructive boundaries

When an ocean plate subducts beneath another plate, it rubs against the plate above it and gets hot. The rock melts resulting in a magma chamber under the upper plate. This magma forces its way through weak points in the crust. This creates a line of volcanoes parallel to the boundary but off to one side in the upper plate. Most of the world's volcanoes occur at destructive boundaries. The ring of fire in the Pacific Ocean is a good example of volcanoes along destructive boundaries.

Volcanoes at constructive boundaries

Volcanoes can also form along the constructive boundaries. These are formed where two plates move apart at the boundary and the magma moves up from the underneath magma chamber to fill the gap left by the separating plates. Ol Doinyo Lengai mountain in Tanzania is an example of a volcano at a constructive boundary. Also the Mid-Atlantic Ridge that runs along the floor of the Atlantic Ocean is an example of a volcano formed at a constructive boundary.

Hot spot volcanoes

Volcanic eruption may also occur thousands of kilometres away from the tectonic plate boundaries. It is thought that these eruptions occur over places in the mantle that are hotter than normal. Magma from these hot spots forces its way through the crust above and onto the Earth's surface. This process is illustrated in Figure 7.7. Nyamuragira volcano in Congo is an example of a hotspot volcano.

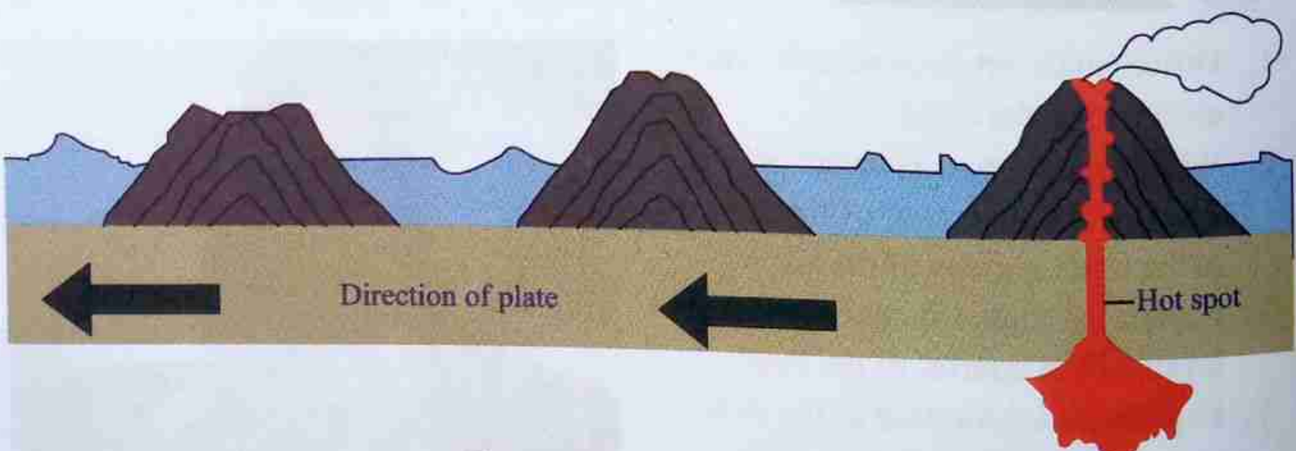


Figure 7.7: *Hot spot volcano*

Types of volcanoes

Volcanoes are categorized based on the mechanism of lava eruption and on frequency of eruption. Based on the mechanism of lava eruption, there are two types of volcanoes, namely *central* and *fissure* volcanoes. In central volcanoes lava erupts to the Earth's surface through a single vertical main vent, whereas in fissure volcanoes, lava erupts through multiple elongated cracks or fractures on the Earth's surface. The length of these cracks can be many kilometres. In all volcanoes, if gases and rock fragments are released to the atmosphere through the vent, the eruption is termed as *explosive* eruption. Conversely, if during eruption lava flows slowly, the eruption is called *effusive* eruption.

On the other hand, based on the frequency of eruption, volcanoes are classified into three categories namely *active*, *dormant* and *extinct* volcanoes.

Active volcanoes are those that either erupt constantly or have erupted in recent times. Good examples of such volcanoes include the Mount Ol Doinyo Lengai in Tanzania, Mount Longonot in Kenya and Mount Nyiragongo in the Democratic Republic of Congo.

Dormant volcanoes are the volcanoes that have been inactive for some time (a few thousand years) but are capable of erupting again. Examples of dormant volcanoes are Mount Kilimanjaro and Mount Meru in Tanzania.

Extinct volcanoes are volcanoes that erupted previously but have not erupted in the recorded history of at least 10 000 years. Such volcanoes may probably never erupt again. Examples of extinct volcano are Izumbwe - Mpoli Mountain in Tanzania, mount Elgon at Kenya-Uganda boundary and the Ben Nevis Mountain in the United Kingdom.

Effects of volcanic eruption

Volcanoes can have both positive and negative effects. The following are some of the effects of volcanoes.

Positive effects

Geothermal energy – In areas where magma is close to the surface, geothermal energy can be harnessed.

Soil fertility – Volcanoes help in soil formation and increase soil fertility by bringing important soil minerals from deep underground onto the Earth's surface.

Minerals – Volcanic eruptions also bring valuable minerals to the Earth's surface such as copper, zinc, silver and gold. These minerals are important economic resources.

Negative effects

Loss of vegetation and wildlife – Volcanic eruptions destroy vegetation as the eruptions sometimes set the surrounding vegetation on fire. Such fires consume huge tracts of vegetation including forests, woodlands and grasslands. Wild animals and other species are also killed by being buried in the lava or being burnt by the forest fires. This leads to loss of biodiversity.

Loss of human life and property – Volcanic eruptions sometimes kill people and destroy property. People who monitor volcanic activities usually warn people of an impending eruption so that they can vacate areas that are susceptible to the volcanic eruption. However, some eruptions happen unexpectedly. Such eruptions kill people, animals and destroy infrastructure around the volcano. In this case it is not possible to give forewarning for people to be evacuated.

Environmental pollution – Volcanic eruptions emit large amount of harmful gases into the environment such as sulfur dioxide, carbon dioxide, methane, argon, carbon monoxide and water vapour. Some of these gases contribute to global warming, climate change and pose a threat to ecosystems.

Damaged landscape - Lava flow alters the original morphology of the land in the area around the volcano. Structures like volcanic mountains, craters are results of volcano eruption. Figure 7.8 illustrates the damage done by volcanic eruption.

Figure 7.8: *Effect of volcanic eruption*

Task 7.2

In groups, discuss how density of magma affects lava flow. Present your findings in the class.

Earthquakes

Earthquakes occur when rocks in the Earth's crust move suddenly, shaking the Earth. About 10 000 earthquakes occur every year but most of them are so small that they can only be detected by very sensitive instruments. Earthquakes also occur as a result of movement of magma at constructive boundaries, under volcanoes and where continental plates collide and push mountain ranges. They range in size from those having undetectable magnitudes to those having enough energy to shake and displace or disrupt the ground.

The origin of earthquakes

Earthquakes mostly occur on or near the boundaries of tectonic plates. However, they can also occur far from the plate boundaries. Earthquakes that occur far from a plate boundary are probably a result of faults that were formed millions of years ago.

Earthquakes that occur on or near a plate boundary are triggered by a sudden release of energy inside a small area of the Earth's rocks. Tectonic plates grind past each other, rather than slide past each other smoothly. As the plates move past each other they can become locked together due to friction. For some time, these plates do not move and stored elastic energy builds up. Figure 7.9 shows two plates that are momentarily held by friction between them.

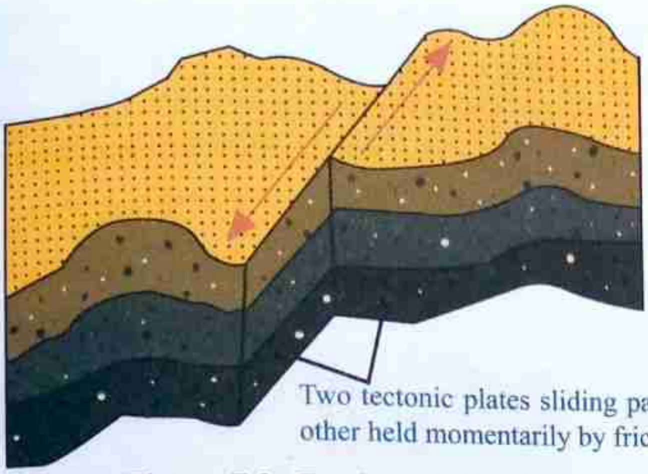


Figure 7.9: *Earthquake build-up*

Pressure builds between tectonic plates until the frictional force holding the plates together is overcome. Consequently, the plates move suddenly, releasing the energy and then hold together again. This results to sudden shaking of the Earth surface. This shaking is what is felt as the *earthquake*.

An earthquake is a sudden release of energy that has accumulated within or along the edges of the earth's tectonic plates, resulting in sudden motion or shaking of the earth.

The point within the earth where an earthquake originates is called the *hypocentre* (*or focus*) of the earthquake. The point on the surface of the earth and directly above the focus is known as the *epicentre*. Figure 7.10 shows the hypocentre and epicentre of earthquake.

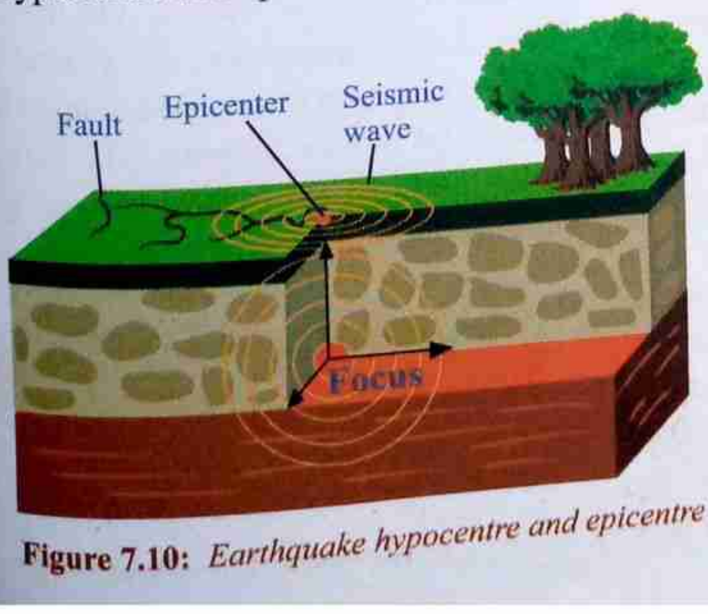


Figure 7.10: *Earthquake hypocentre and epicentre*

Seismic waves

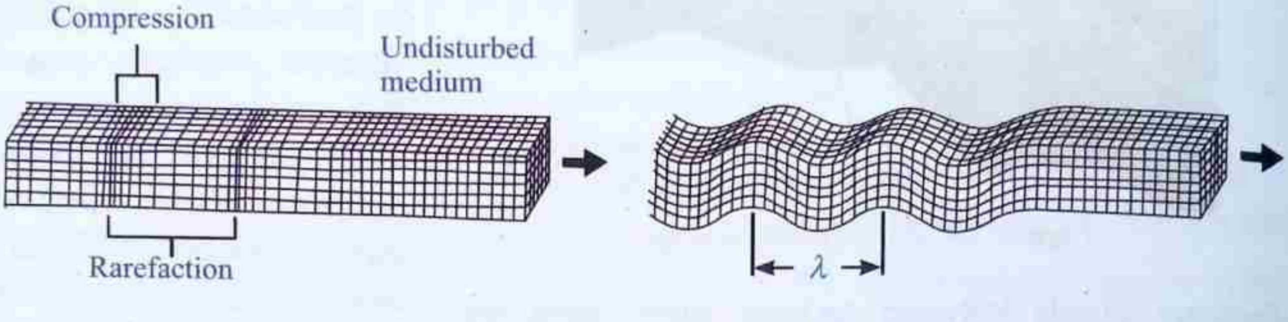
The energy released during an earthquake is transmitted in form of waves called *seismic waves*. Seismic waves can be generated by earthquakes, volcanic eruptions, magma movements, large landslides and large man-made explosions. These waves can travel through the Earth's interior or along its surface. The propagation velocity of seismic waves is influenced by density and elasticity of the medium as well as the mode of propagation. There are two broad types of seismic waves. These are *body waves* and *surface waves*.

Body waves

These are seismic waves which travel through the interior of the Earth. These waves are further divided into *Primary waves (P-wave)* and *Secondary waves (S- waves)*.

Primary waves (P-waves) are compressional waves that travel longitudinally as displayed in Figure 7.11 (a). They travel quicker through the Earth than other waves and they are the first to reach a seismograph station, hence the term "Primary". They can move through solids and liquids.

Secondary waves (S-waves) are transverse in nature, and can only move through solids and water surface (Figure 7.11 (b)). They arrive at a seismograph station after the P-waves and dislocate the ground perpendicular to the direction of propagation. They are felt as a series of side-to-side tremors. Table 7.1 shows the difference between S-wave and P-wave.



(a) Primary waves (b) Secondary waves

Figure 7.11: *Propagation of body waves*

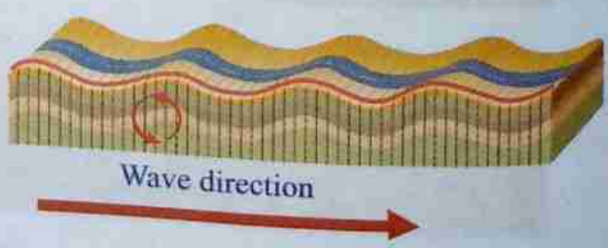
Table 7.1: *Difference between S-wave and P-wave*

P- wave	S-waves
Longitudinal (compressional)	Transverse waves
Travel at a speed of 7 - 14 km s^{-1}	Travel at a speed of 3.5 - 7 km s^{-1}
Can propagate in solid and liquid, hence can pass through crust, mantle and core	Cannot propagate in fluids (liquids and gases), its velocity depends on density and shear modulus of medium material
Reach the earthquake recording station before S-waves	Reach the earthquake recording station after the P-waves

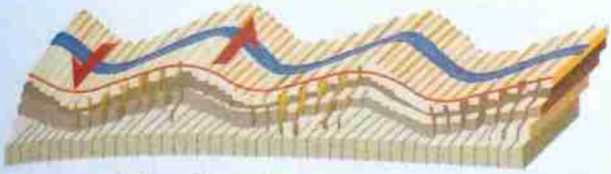
Surface waves

These are mechanical waves that move along the surface of the Earth. That is, they propagate only at the interface between two different media. The amplitude of surface waves get smaller as one moves far from the surface. They move at a slower speed than the body waves.

There are two types of surface waves; these are *Rayleigh waves* and *Love waves*. Rayleigh waves create a rolling movement that makes the land surface move up and down and forward and backward in the direction of the wave (Figure 7.12 (a)). Their amplitude decreases dramatically with depth. Most of the shaking of the earth is caused by these waves. On the other hand, Love waves make the ground shift from side to side hence producing entirely horizontal motion (Figure 7.12 (b)). The amplitude of Love waves is large at the surface and diminishes with depth.



(a) Rayleight waves



(b) Love waves

Figure 7.12: *Surface waves*

It is the surface waves that damage the surface structures such as buildings and hydroelectric power plants. Figure 7.13 shows a damage that has been caused by an earthquake.

Figure 7.13: *Effect of an earthquake*

Principles of measurement of earthquakes

In order to fully understand the nature of earthquakes, two properties are measured. These are the *magnitude* and *intensity*. The magnitude of an earthquake is a measure of the energy it releases. It is usually measured on the *Richter scale*. The Richter scale is based on the magnitude of a recorded seismic wave of an earthquake, without considering which type of wave was the strongest. The magnitudes on the Richter scale are based on a logarithmic scale (base 10). This specifies that the amplitude of the ground motion increases ten times for every whole number on the Richter scale. An earthquake with a magnitude of 7.0 would cause 10 times the amount of ground shaking as an earthquake with a magnitude of 6.0. The Richter scale can even be used to describe earthquakes that are so minor that the values are negative. There is no limit to how high the scale can go.

On the other hand, the intensity of an earthquake is a measure of its severity based on the effects it causes to the landscape. The intensity of an earthquake is usually measured on the *Modified Mercalli scale*. The scale is calibrated in levels from I to XII, for which level I is a minor shock that causes no damage whereas level XII causes total damage. Table 7.2 shows a description of the twelve levels on the Modified Mercalli scale.

Table 7.2: *Twelve levels on the Modified Mercalli scale.*

Intensity level	Shaking	Description / Effects
I	Not felt	It is not felt, except by a small number of people under unusually favourable circumstances. Magneto receptive species can also sense the earthquake at low intensity. Examples are rats, bats and chicken.
II	Weak	Only a few people, especially on upper floors of buildings, can feel it when they are at rest. Objects that are precisely suspended may swing.

III	Light	People indoors, especially on higher floors of buildings, can feel it fairly strongly. It is not easily recognized by many people. It is possible that standing automobiles will shake a little. Vibrations resemble a passing truck. It is possible to predict the duration.
IV	Moderate	During the day, many people feel it indoors, while only a few people feel it outside. Some people awoke in the middle of the night. Dishes, windows, and doors may be disturbed. The sound of cracking walls can be heard. A sensation similar to that of a large truck colliding with a structure. Cars that were parked are disturbed a lot.
V	Rather strong	Almost everyone may feel it. Many people are awakened. Some dishes and the windows are broken. Unstable things tipped over.
VI	Strong	Everyone can sense it. Many people are terrified. Some large pieces of furniture are moved. The damage is minor.
VII	Very strong	Damage is negligible in buildings of good design and construction. Slight to moderate damage in well-built ordinary structures. Poorly constructed buildings suffer extensive harm. It is possible that some walls will collapse.
VIII	Severe	In well-built structures, the damage is minor. Ordinary buildings may have partial damage. In poorly constructed structures, the amount of damage is significant. Chimneys, factory stacks, columns, monuments, and walls may fall to the ground.
IX	Violent	In special built structures, the damage is significant. Frame constructions that were well-designed may fall. Significant damage to large buildings, with partial collapse. Buildings begin to shift from their foundations.
X	Intense	Some well-constructed wooden structures are demolished. The rails are deformed.
XI	Extreme	Few masonry structures remain standing. Bridges demolished. Rails bent significantly.
XII	Catastrophic	Total devastation. The level and lines of sight are altered. Objects are hurled across the air.

Note that an earthquake can only have one magnitude; the value refers to the magnitude at the focus. However, its intensity reduces as the seismic waves spread out from the hypocentre, just the same way the loudness of sound changes as you move away from the source.

The seismograph

The vibrations within and at the surface of the Earth can be detected using a *seismograph*. A seismograph measures ground oscillations by recording the relative motion between a pendulum and the ground. It is also possible to use the ratio between the deflection of the pendulum and the acceleration of the ground to record an earthquake. The time of initiation of ground oscillations is recorded and marked on the graphs every minute and hour on the seismograph paper. Figure 7.14 shows a seismograph.

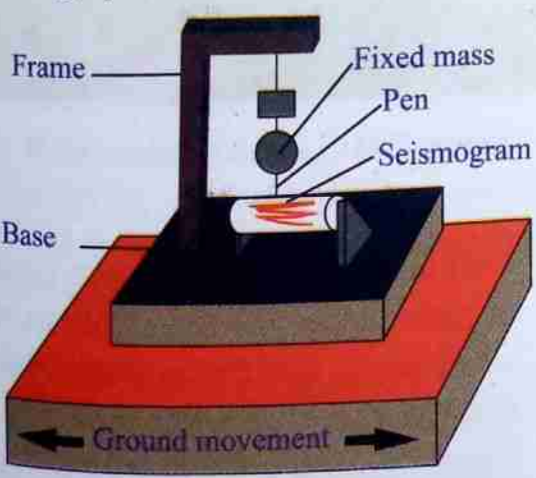
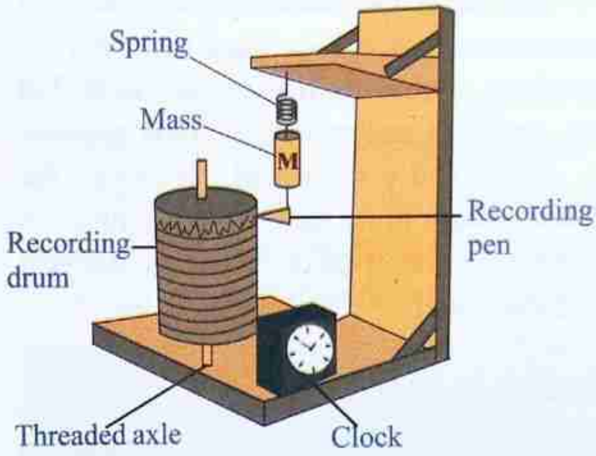
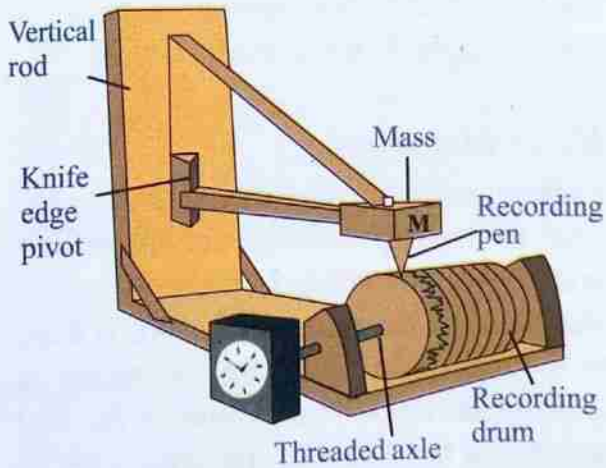


Figure 7.14: *Seismograph*

In order to measure ground motions, the seismograph must remain steady when the ground moves. Various types of pendulums have been used for this purpose. The simplest type of pendulum is a heavy mass suspended by a wire or rod from a fixed point. Other forms are the inverted (vertical) and horizontal pendulums. The inverted pendulum has a heavy mass fixed to the upper end of a vertical rod and pointed at its lower end (Figure 7.15 (a)), while the horizontal pendulum has a rod with a mass on its end which is suspended at two points so that it swings on a horizontal plane as shown in Figure 7.15 (b).



(a) Vertical pendulum seismograph



(b) Horizontal pendulum seismograph

Figure 7.15: *Inverted/vertical and horizontal pendulum seismograph*

Recording the pendulum motion

The recording of the motion of the pendulum can be done in various ways. The most common ones are the mechanical method, the optical method and the electronic method.

Mechanical method

In the mechanical method, a sheet of smoked paper is wrapped round a clock-driven rotating drum and mounted to move with the Earth. A moving pen connected to the pendulum presses lightly on the paper. As time passes, the drum rotates so that the recorded lines are not superimposed on each other. Deflection of the pendulum is commonly magnified mechanically by single or double multiplying levers so that the graph is easier to see. This method is simple and economical. However, the seismograph must have a heavy mass to overcome the friction between the pen and the paper.

Optical method

The optical method still uses a pendulum motion to record the ground movements. However, to overcome friction, mirrors are used to reflect the light onto photosensitive paper wrapped on the drum.

Electronic method

Technological developments have given rise to high-precision seismographs and sensors of ground motion. In these electromagnetic instruments, a coil is fixed to the mass of a pendulum and moves in a magnetic field. The motion of the coil generates the e.m.f in the same way as dynamo does. The voltages produced by motions of the pendulum are passed through electronic circuits for amplification in order to get reliable readings.

Hazards of Earthquake

Earthquakes give rise to a number of hazards which pose a great risk to human life, animals, property and the environment. The following are some of the hazards associated with earthquakes

1. **Landslides**– The shaking caused by earthquakes can cause unstable hillsides, mountain slopes and cliffs to move downwards, creating landslides. In massive landslides soil and rocks accelerate down slopes, sweeping away everything in their path. Landslides can block valleys and streams channels, creating temporary dams. These dams release water when they collapse leading to floods. Earthquakes can also trigger avalanches on snow slopes. Figure 7.16 shows a landslide caused by earthquake.

Figure 7.16: *Landslide caused by an earthquake*

2. **Tsunamis**– If an earthquake occurs in rocks under the sea or ocean, the shock waves disturb the water. This is because, as the ocean floor rise or fall it causes the water to rise or fall too. These movements create huge water waves called tsunamis that travel across the ocean. When a tsunami reaches shallow water, it slows down, its wavelength reduces and its height grows. When a tsunami hits the shore, it crashes the land carrying everything in its way including buildings. Figure 7.17 shows a tsunami approaching a city.

Figure 7.17: *Effects of tsunami*

3. **Collapsing buildings**- Earthquakes do not actually kill people. It is the effects associated with earthquakes that kill people. The majority of people killed or injured in earthquakes are trapped in buildings that collapse due to the shaking of the ground underneath. A strong earthquake can flatten a whole city. An example is the Japanese city of Kobe shown in Figure 7.18. The city was completely flattened by an earthquake that occurred in January 1995, measuring 7.2 on the Richter scale. More than 6 000 people died and about 200000 buildings collapsed or were damaged.

Figure 7.18: *Effect of earthquake at Kobe city*

4. **Fire outbreak**– An earthquake can trigger a fire outbreak. This happens when the earthquake causes gas or oil pipes to break. It can also occur as a result of the collapse of electricity lines.

5. **Backward rivers**- Tilting ground can also make rivers change their course. This can result in the creation of earthquake lakes that cover huge tracts of previously settled land. Figure 7.19 shows a backward river formed as a result of an earthquake.

Figure 7.19: *Backward river*

Possible indicators for earthquake occurrence

The following are indicators of the possibility of occurance of an earthquake.

1. **Thermal indicator**- A few months before the occurrence of an earthquake, the average temperature of the area keeps increasing. On the day of the earthquake, the temperature of a place is about 5 to 9 degrees Celsius above the average normal temperature for that day.

2. **Water indicator**– About one or three days before an earthquake, there is a sudden rise or fall in water levels in wells. The rise can be as high as one metre. The well water may turn muddy. Sometimes a fountain may appear in the ground. This normally happens a few hours before the earthquake. There is also a sudden and rapid increase or decrease of water flow in the rivers. This happens about one to two days before the earthquake.
3. **Seismo-electromagnetic indicator**– Before the occurrence of an earthquake the sub-surface temperature rises. As a result of this, the geomagnetic field is reduced. The reduction in geomagnetic field adversely affects the propagation of electromagnetic waves. This is experienced abundantly on the radio, television and telephone. This is a very reliable indicator. It is usually recorded about 10 to 20 hours before the earthquake.

   If a particular radio station is received at a frequency of 1 000 kHz, the same station will be received in the potential epicenter area at higher frequencies, about 10 to 30 hours before the earthquake. Similarly, reception of television signals is affected. The mobile phone is one of the most reliable indicators of an impending earthquake. About 100 to 150 minutes before the occurrence of an earthquake, mobile phones stop functioning or malfunctions. Note that all the mentioned indicators are valid only when seen and manifested extensively. Failure of one or two instruments should not be taken as an earthquake indicator.

4. **Animal indicator**– Between 10 and 20 hours before the occurrence of an earthquake, some animals become highly disturbed and restless. They move in a random manner and in fear. Some animals such as dogs, bats, and aves may sense the weak intensity of earthquake before human beings can feel it.

Precautions to be taken during an earthquake

The following are some precautions that can be taken to minimise injuries or death of human beings in an event of an earthquake.

(a) If you are indoors during an earthquake, get under a desk, table or bench. Hold on to one of the legs of the desk or table and cover your head and neck. If there is no table or desk nearby, sit down against an interior wall. An interior wall is less likely to collapse than a wall on the outside of the building.

(b) Pick a safe place where things will not fall on you – away from windows or tall and heavy furniture.

(c) Do not run outside when an earthquake occurs because bricks, roofing and other materials may fall from buildings during and immediately after an earthquake.

(d) Wait in your safe place until the shaking stops, then check to see if you are hurt.

You will be able to help others if you take care of yourself first.

(e) Move carefully and watch out for objects that have fallen or broken, creating hazards. Be ready for additional earthquakes called aftershocks.

(f) Be on the lookout for fires. Fire is the most common earthquake-related hazard due to damaged gas and electrical lines.

(g) If you must leave a building after the shaking stops, use the stairs and not the elevator. Earthquakes can cause fire alarms and fire sprinklers to go off. You will not be certain whether there is a real threat of fire. As a precaution, use the stairs.

(h) If you are outside during an earthquake, stay outside. Move away from buildings, trees, streetlights and power lines. Remember to cover your head as you might be hit by objects.

Exercise 7.2

1. Why are some volcanoes called active although they are not erupting?
2. Explain how an earthquake occurs.
3. Does the magnitude and intensity of earthquakes mean the same or not? Explain.
4. What hazards are associated with earthquakes?
5. Describe two types of seismic waves.
6. Explain various ways of recording the motion of the pendulum in a seismograph.
7. A seismometer located in Dodoma recorded an earthquake which originated in Mbeya. Which type of wave was recorded first? Give reasons for your choice.

Structure and composition of the atmosphere

The environment that humans live in is surrounded by gaseous matter around the Earth surface. This environment is called the *atmosphere*. The atmosphere is a mixture of gases and tiny particles and it extends from the Earth's surface outward into space. The two major gases in the atmosphere are nitrogen (78%) and oxygen (21%). Other gases such as carbon dioxide, neon, helium, methane, hydrogen as well as water vapour make up the rest of the atmosphere. Sometimes tiny particles (termed aerosols) are present in the atmosphere. These include particles of sea-salt, mineral dust (particularly silicates), organic matter and smoke.

Structure of the atmosphere

The atmosphere is categorized according to its thermal structure which determines its dynamic properties. Averaging out the temperature values in the atmosphere results into identification of four layers of the atmosphere. These layers are the troposphere, stratosphere, mesosphere and thermosphere. The arrangement of these layers is shown in Figure 7.20.

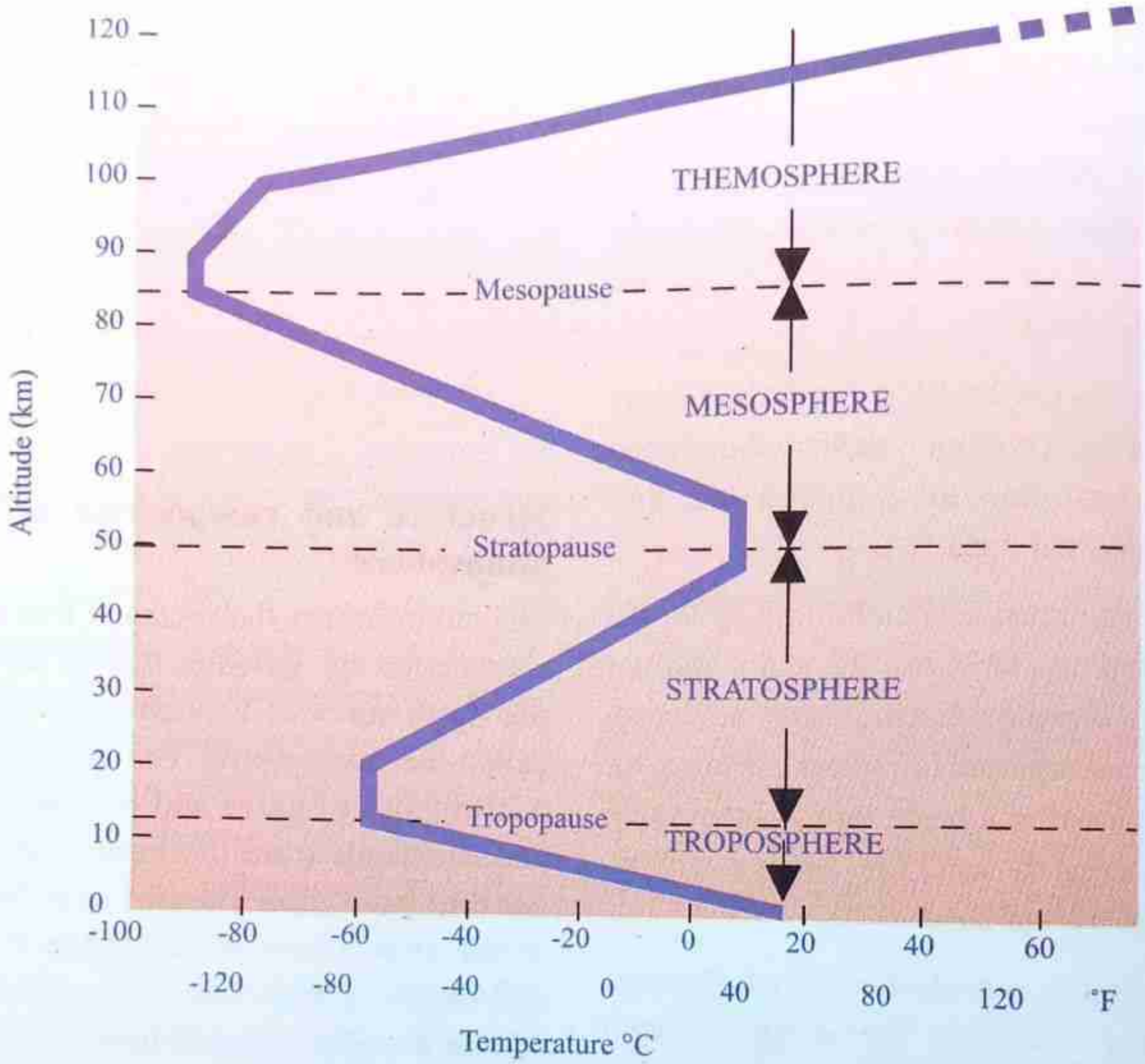


Figure 7.20: *Vertical structure of the atmosphere*

Troposphere

This is the lowest layer which is in contact with the Earth's surface. It is also known as the turning sphere or mixing layer. It extends to an altitude of about 10 km above the poles and about 15 km above the equator. This region is the densest part of the atmosphere, containing about 80% of the mass of the atmosphere. Water vapour and small particles are contained in this layer. The temperature in this region decreases with altitude at an average rate of 6.5 °C/km. The reason for this decrease is that the gases in the troposphere do not absorb solar radiation that enters the atmosphere. Instead, this radiation is absorbed by the ground and then heats the tropospheric air through conduction and convection from below.

The troposphere is well mixed allowing air molecules to travel to the top of the troposphere and return down in just a few days. This mixing encourages changing weather. Most weather phenomena occur in the troposphere. Clouds and rain are formed within this layer. The main components of the troposphere include nitrogen, oxygen, argon, carbon dioxide, water vapour, and methane. Carbon dioxide, water vapour, and methane are commonly known as greenhouse gases. The boundary, a transition region which separates the troposphere and the stratosphere is called the *tropopause*. At the tropopause, the temperature stops decreasing with height and becomes constant.

Stratosphere

The stratosphere starts from the tropopause and extends to about 50 km high. This layer is more stable, drier and less dense compared to the troposphere. It is also called the layered sphere. About 99.9% of the mass of the atmosphere lie below the top of the stratosphere. The temperature in the stratosphere slowly increases with altitude due to the presence of the ozone layer which absorbs ultraviolet rays from the Sun. The ozone layer lies in the middle of the stratosphere between 20 and 30 km of altitudes. This layer plays the important role of absorbing ultraviolet radiations which would otherwise reach the Earth's surface. Ultraviolet radiation is harmful to both animal and plant life on Earth.

The stable air of the stratosphere also prevents large storms from extending much beyond the tropopause. Planes fly in the stratosphere, because it is very dry, free of turbulence and it contains little water vapour. The troposphere and stratosphere are collectively known as the *lower atmosphere*. The boundary which separates the stratosphere and the mesosphere is called *stratopause*.

Mesosphere

The mesosphere starts just above the stratosphere and extends to about 85 km high. The temperature in this layer decreases with altitude because there are few gas molecules to absorb the Sun's radiation. (The main source of heating is the straposphere). The mesosphere is very important for Earth's protection since most of meteors and asteroids are burnt up here before they reach the Earth's surface. The boundary which separates the mesosphere and the thermosphere is called the *mesopause*.

Thermosphere

This layer starts just above the mesosphere and extends up to 690 km high. The temperature increases with increasing altitude due to absorption of highly energetic solar radiation by oxygen molecules. The temperature in this region can go as high as 1 727 °C. Chemical reactions occur much faster here than on the surface of the Earth. This layer is also known as the *upper atmosphere*.

The lower part of the thermosphere, from about 80 to 550 km above the Earth's surface, contains the *ionosphere*. This is a region containing a high concentration of charged particles called ions and free electrons. The large number of free electrons in the ionosphere allows the propagation of electromagnetic waves. Radio waves can be reflected off the ionosphere allowing radio communications over long distances. More ions formation leads to more empty space, the *exosphere*.

Importance of the atmosphere

The following are some ways in which the layers of the atmosphere are important:

(a) The troposphere controls the climate and ultimately determines the quality of life on the Earth.

(b) The troposphere contains gases including oxygen which is used for respiration by animals and carbon dioxide which is used by plants in photosynthesis. The nitrogen found in this layer also provides an inactive environment for many chemical processes to take place. The gases also support many important processes such as combustion, weathering and oxidation.

(c) The stratosphere prevents harmful ultraviolet radiation from reaching the Earth as they are absorbed by the ozone.

(d) The mesosphere and thermosphere prevent harmful radiation such as cosmic rays from reaching the Earth's surface.

(e) Communication is also made possible by some layers of the atmosphere, specifically the ionosphere.

(f) Most of meteors and asteroids are burnt up in the mesosphere before they reach the Earth's surface.

Exercise 7.3

1. Describe the composition of the atmosphere.
2. Briefly explain the importance of the stratosphere to living things on the Earth's surface.
3. What is the importance of the large number of free electrons in the ionosphere?
4. Why is it important for a captain of an aircraft to know the height of the tropopause?
5. Visualize temperature as a function of altitude from 0 to 100 km as shown in Figure 7.20. Why is the trend not the same in all the layers of the atmosphere?

The greenhouse effect and global warming

Scientists investigate the link between greenhouse gases in the atmosphere and global warming. They are trying to figure out how increased greenhouse gas emissions may contribute or have contributed to global warming. It is therefore important to learn about the greenhouse effect.

The greenhouse effect

Greenhouse effect is a process in which gases in the atmosphere trap thermal radiation that is emitted from the Earth's surface and re-radiates it back to the Earth's surface.

The Sun provides energy in the form of ultraviolet, visible, and near-infrared radiation to the Earth. The atmosphere reflects some of the incoming solar energy back to space, and absorb some of it. The majority of the remaining energy is absorbed at the Earth's surface. The Earth then re-radiates at considerably longer wavelengths than those that were absorbed. Long wave radiation (in the infra-red region) emitted from the Earth's surface is trapped by the greenhouse gas molecules in the lower Earth's atmosphere. This radiation cannot escape to space. The excited molecules re-emit the absorbed energy which is also a long wave radiation back to the Earth's surface. The net effect is trapping of the heat in the atmosphere and hence warming up the Earth's surface to a comfortable level. Before the industrial revolution, concentrations of greenhouse gases, especially Carbon dioxide, in the atmosphere were at their natural levels. Hence the natural greenhouse effect was responsible for maintaining the Earth's temperature at a comfortable natural level. Figure 7.21 illustrates the greenhouse effect.

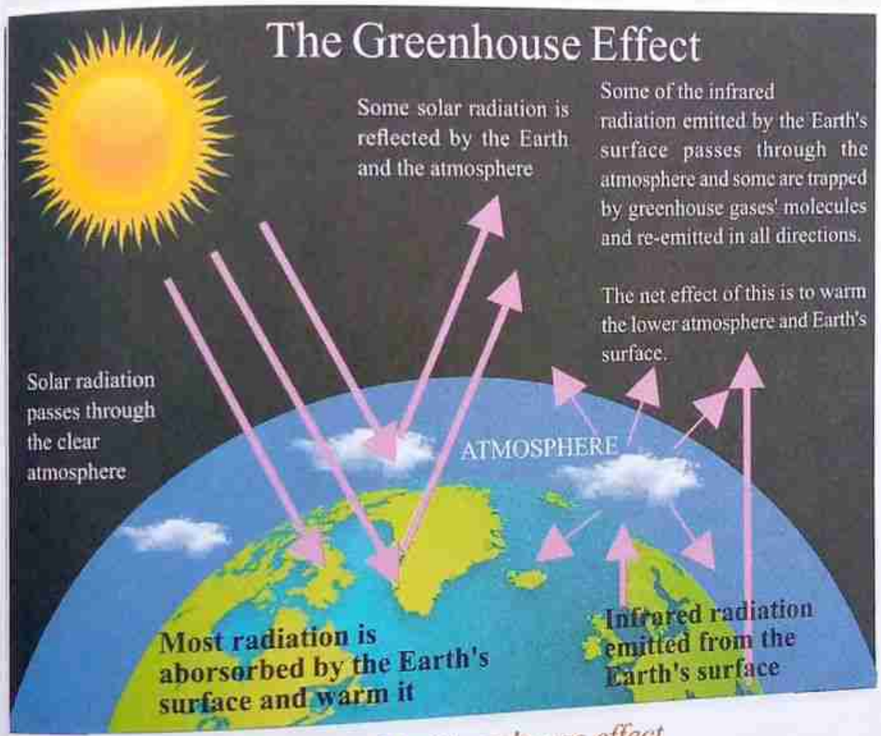


Figure 7.21: *Greenhouse effect*

Human activities have forced the increase of greenhouse gases in the atmosphere. These gases absorb and re-emit long wave radiation to the Earth surface resulting to the increase of the Earth's temperature. This is called enhanced greenhouse effect and is responsible for the global increase of the average temperature near the surface.

Sources of greenhouse gases

Gases in the atmosphere that absorb and emit infrared radiation in the wavelength range emitted by Earth's surface are called greenhouse gases. These gases are responsible for greenhouse effect and they are important in balancing the Earth's temperature. The main greenhouse gases include: carbon dioxide; methane; water vapour; and nitrous oxide.

Carbon dioxide

Carbon dioxide is the main greenhouse gas. The gas contributes to over 50% of the greenhouse effect. The following are some of the sources of carbon dioxide in the atmosphere:

(a) ***Deforestation and burning of vegetation***

Green plants, especially woody plants, absorb carbon dioxide from the atmosphere as they grow. Clearing of forests (deforestation) and burning vegetation results in the release of carbon dioxide to the atmosphere. The loss of the forests also means that there are fewer trees to absorb carbon dioxide. Figure 7.22 shows a deforested land.

The rate of deforestation remains high in many parts of the world.

Figure 7.22: *Deforestation contributes to global warming*

(b) *Burning of fossil fuels*

Carbon dioxide is a by-product in complete combustion of fossil fuels such as coal, gasoline and natural gas. These fossil fuels are burnt in cars, power stations and industries.

Methane

Methane is emitted from both natural sources and human activities. Natural sources of methane are wetlands, termites, and the oceans. Human based source of methane includes agricultural activities, fossil fuel production and intensive livestock farming. Other sources are landfills and waste, biomass burning, as well as burning biofuels such as ethanol. Due to these sources, the amount of methane in the atmosphere has been increasing by 1% annually since 1960s. This rate is twice the rate of increase of Carbon dioxide in the atmosphere.

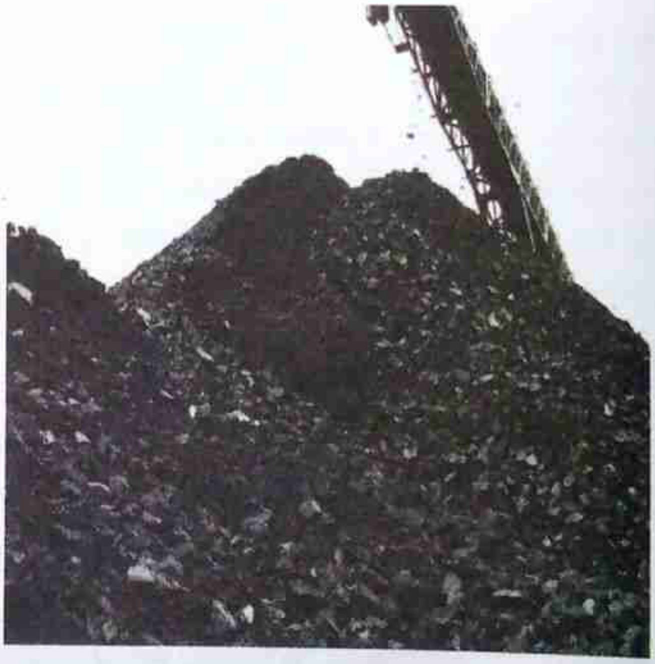


Figure 7.23: *Coal mining station*

Nitrous oxide

Nitrous oxide is a by-product of fuel burning. When any fossil fuel is burned, a portion of the nitrogen in the fuel as well as the surrounding air is oxidized, resulting in nitrous oxide emissions. The majority of stationary emissions come from coal fired power plants. As for mobile emissions, almost all of it comes from cars and trucks. Likewise, the burning of vegetation and animal waste contribute to the emission of nitrous oxide.

Task 7.3

In groups, discuss the possible sources of greenhouse gases around your locality. How can you control them? Present your results in class.

Occurrence of global warming

Global warming is the increase of average global temperature near or on the surface of the Earth as a result of the enhanced greenhouse effect. The rapid increase of greenhouse gases has raised the concern on global warming. Naturally, the amount of radiation (energy) entering the Earth-atmosphere system must be equal to radiation leaving the system, otherwise the Earth surface would continually heat or cool. At natural concentration of greenhouse gases, the energy entering the Earth surface balances that leaving the Earth surface. However, as the concentration of greenhouse gases is increasing due to human activities, the energy leaving the Earth surface is less than that entering the surface. This is because the increased greenhouse gases absorb more long-wave radiation and re-emit back to the Earth's surface resulting to more heating on the Earth's surface and hence global warming.

Effects of global warming

One of the most visible and evident results of global warming is the increase in temperatures around the world. Over the last 100 years, the average global temperature has risen by around 0.8 degrees Celsius. The following are the consequences of global warming:

On biodiversity

Global warming impacts biodiversity. Climate change and rising temperatures disrupt ecosystems, altering plant reproductive conditions and cycles. The habitats of animals and migratory cycles are affected by resource constraint. Many species, especially indigenous species, can be extinct or invasive species can invade and pose a hazard to crops and other creatures.

On oceans

Sea levels are rising due to thermal expansion of the oceans, melting of ice sheets and glaciers. This may eventually lead to flooding of coastal lands. The world's oceans soak up much of the carbon dioxide produced by living organism either in the form of dissolved gas, or from the skeletons of tiny marine creatures that fall to the ocean bottom to become limestone. The carbon dioxide dissolves in the water and forms a weak carbonic acid, thereby lowering the pH of the ocean waters. Increased acidity and temperatures of ocean waters eventually lead to bleaching and death of coral reefs.

On climate

The climatic pattern in most parts of the world has changed. It is becoming hard to forecast the weather accurately. Rain no longer falls when expected. Sometimes the rains are heavier than expected, leading to flooding and decreasing water quality. Other times, the rains are far less than expected, leading to drought. The extent of the Earth's surface under desert condition is also increasing. Occurance of other extreme weather events includes heat waves, hurricanes (Figure 7.24) and tornadoes.

Figure 7.24: *Hurricane winds*

Solutions to global warming

The effects of greenhouse gases in the atmosphere will continue to be felt for a long time. This is because greenhouse gases remain in the atmosphere for long periods of time. For example, carbon dioxide molecules have a lifetime ranging between 50 and 100 years in the atmosphere. This means that global warming will continue even if we were to cut down on the emission of the greenhouse gases.

However, it is important to reduce the amount of the emissions of greenhouse gases. The following are some of the measures that can be taken to reduce the emission of greenhouse gases into the atmosphere:

(i) Put in place energy-conservation measures to reduce the use of fossil fuels. These measures include use of public transport to minimize the number of vehicles on the roads and the use of fuel-efficient cars.

(ii) Use of cleaner alternative sources of energy, such as solar energy, wind energy, biomass and geothermal.

(iii) Replant trees (afforestation) that would absorb carbon dioxide and avoid massive deforestation. Figure 7.25 shows an area with trees necessary for controlling global warming.

Figure 7.25: *Planting trees can help control global warming*

(iv) Recycling is an absolute necessity. Countries should set policies that encourage recycling of waste products. Typical materials that can be recycled include iron and steel scrap, aluminium cans, glass bottles, paper, wood and plastics.

(v) Emissions of carbondioxide from buildings should be minimized by designing buildings that will require less heating, cooling or lighting.

(vi) Government to adopt carbon emissions policies. Carbon pricing policies require companies to pay a price for each tone of released carbon. A carbon tax be imposed on business and individual that works as a sort of pollution tax.

(vii) Farmers should be encouraged to invest on organic farming, which is free from fertilizers that adds pollutants which contribute to global warming.

Exercise 7.4

1. Explain the concept of greenhouse effect.
2. In your village, people are complaining on the increase of temperature nowadays.
   (a) What are you going to explain to them about the possible causes of the prevailing temperature?
   (b) Clarify to them, the measures to be taken to mitigate the situation.
3. How can planting trees reduce the amount of carbon dioxide in the atmosphere?
4. You are working at a village where people are mainly using charcoal as a source of energy. Explain to them the effects that might arise to the environment as a result of prolonged use of charcoal.
5. Explain the side effects of industrial development with respect to global warming.
6. Explain why radiation originating from the Earth is termed as long wave radiation.

Chapter summary

1. The interior of the Earth is composed of three main layers, namely the crust, mantle and core.
2. The Earth's crust and part of the mantle are cracked into huge pieces called tectonic plates which move at very low speeds.
3. There are three main types of tectonic plate boundaries. These are destructive boundaries, constructive boundaries and conservative boundaries.
4. Volcanoes are classified into central and fissure volcanoes based on the mechanism of lava eruption. They can also be categorized into active, dormant and extinct volcanoes based on the frequency of eruption.
5. Earthquakes occur when rocks within the Earth's crust move suddenly, shaking the Earth.
6. The point within the Earth where an earthquake begins is called the hypocentre or the focus of the earthquake, and the point on the Earth's surface above the hypocentre is called the epicentre.
7. Earthquakes release their energy in form of waves called seismic waves. Seismic waves can be body waves comprising of primary waves and secondary waves, or surface waves

comprising of Rayleigh waves and Love waves.

8. The magnitude of an earthquake is usually measured on the Richter scale while the intensity of an earhquake is measured on the Modified Mercalli scale.
9. A seismograph is an instrument used to record ground movements caused by an earthquake.
10. Hazards associated with earthquakes include landslides, tsunamis, collapsing buildings, fire outbreaks and backward rivers.
11. The Earth's atmosphere is divided into four vertical layers based on temperature distribution. These are the troposphere, stratosphere, mesosphere and the thermosphere.
12. Global warming is the increase in the average temperature on or near the Earth's surface as a result of the enhanced greenhouse effect. It may result to the melting of polar ice, rise in the sea levels, change of climatic patterns, oceans acidification as well as extreme weather events such as heat waves, drought, floods and hurricanes.
13. The greenhouse effect is caused by greenhouse gases, which include carbon dioxide, methane, nitrous oxide and water vapour.

Revision exercise 7

Choose the most correct answer in items 1-4.

1. Which among the following is not categorized under geophysical hazard?
   (a) Tropical cyclone
   (b) Earthquakes
   (c) Volcano
   (d) Floods
2. Which of the following waves is a body wave with low speed?
   (a) P-waves
   (b) S-waves
   (c) Rayleigh waves
   (d) Love waves
3. What layer separates the crust from the core of the Earth?
   (a) Magma layer
   (b) Lithosphere
   (c) Mantle
   (d) Continent
4. With the current increasing global temperature, what is expected to happen to the sea level?
   (a) The sea level will fall.
   (b) The sea level will rise.
   (c) The sea level will remain unchanged.
   (d) The sea level will rise and fall.
5. Draw a cross section of the Earth's interior region and use it to show the:
   (a) Crust.
   (b) Mantle.
   (c) Outer core.
   (d) Inner core.

6. Explain why the inner core of the Earth is solid while the outer core is liquid even though the inner core has a higher temperature than the outer core.

7. What is the difference between crust and lithosphere? Where are they located and what are their properties?

8. How does the difference between oceanic and continental crust lead to the presence of oceanic basins and continents?

9. (a) Explain how volcanoes occur.
   (b) Distinguish between central and fissure volcanoes.

10. (a) What is an earthquake?
    (b) Explain the meaning of the following terms as used on earthquakes:
       (i) Hypocentre.
       (ii) Epicentre.
    (c) Name the instrument used to measure an earthquake.
    (d) List down three precautions that can be taken against earthquake hazards.

11. If energy can be neither created nor destroyed, what happens to the energy released during an earthquake?

12. Can an earthquake cause a volcanic eruption and vice versa? Explain.

13. (a) List down the layers of the atmosphere starting from the Earth's surface.
    (b) Give one importance of each of the layers of the atmosphere.
    (c) What property that makes the ionosphere an important layer? Explain.

14. (a) What is global warming?
    (b) Name four gases that contribute to global warming. For each gas, write at least one source producing it.
    (c) Write at least four effects of global warming.
    (d) Explain at least three measures that can be taken to control global warming.

15. Figure 7.26 shows a certain human activity.

Figure 7.26

(a) Explain how the activity contributes to global warming.
(b) How can the activity be controlled?

Answer to numerical questions

Chapter One

Exercise 1.1

2. 2.50 m
3. 6.60 m s^{-1}
4. 0.70 m s^{-1}
5. 32 cm or 0.32 m
6. (a) 1.10

   (b) 33.37° or 33°22′
7. 5.50 Hz
8. (a) 2 m

   (b) 2×10^8 m s^{-1}
9. 3 cm; 21 cm s^{-1}; 7 Hz
10. 660 Hz

Exercise 1.2

2. 1450 m s^{-1}
3. 330 m s^{-1}
4. 742.50 m
5. 0.29 s
6. 1783:1837
7. (a) 333.33 m s^{-1}

   (b) 100.35 m

Exercise 1.3

1. (b) 1.2 m
2. (a) 353. 55 Hz

   (b) 1060.66 Hz
3. $4f$
4. (a) 1200 Hz

   (b) 800 Hz
5. (a) 344 m s^{-1}

   (b) 3.4 cm
6. 680 Hz
7. 320 Hz

Revision exercise 1

4. (a) 2 m

   (b) 0.004 s

   (c) 250 Hz

   (d) 8 m

   (e) 2×10^3 m s^{-1}
6. 2437.40 m
7. 1378.40 m
8. 25 Hz, 75 Hz
12. 1.41 m
13. 27 cm
15. 6×10^{14} Hz
16. 1.28×10^{-2} s

19. (c) (i) 800 Hz (ii) 4.50 N

20. (a) 58 cm (b) 22.50 N

Chapter Two

Exercise 2.3

1. (a) 20 V
   (b) 96.67%
2. 64 A
3. (a) 1.67 A
   (b) 12 V
   (c) 10:1

Revision exercise 2

10. 1:2
11. (a) 0.25 A (b) 2.5 A
12. (a) 2400 V (b) 50%
15. 20 V
16. 4:1
19. (a) 2.4 V (b) 7.2 V
    (c) 60 (d) 9.6 V
    (e) 1.6 A

Chapter Three

Exercise 3.2

4. (a) 8 (b) 0 C
6. (a) 11 electrons, 13 neutrons
   (b) 1.76×10^{-18} C

Exercise 3.3

1. 10 years
2. 120 000 years
3. 210 g
4. $\frac{1}{8}$

Revision exercise 3

7. (a) 8
   (b) 1.28×10^{-18} C
18. 2.5 mg
19. $\frac{1}{4}$

Chapter Five

Revision exercise 1

17. (c) 400 μA
18. (b) 0.05 mA
    (c) 65

Chapter Six

Revision exercise 6

8. 40 kg, 66.67 N
9. 216N, 0.34 kg m^{-3}
21. (a) 4.76×10^{-6} light years
    (b) 350 s

Appendix I: Periodic table

- Alkali metals
- Alkaline earth metals
- Transition metals
- Poor metals
- Metalloids
- Non-metals
- Noble gases
- Rare earth metals

	Group I	Group II	Transition metals										Group III	Group IV	Group V	Group VI	Group VII	Group VIII/0
Period 1	1 $_{1}$H Hydrogen																	4 $_{2}$He Helium
Period 2	6.9 $_{3}$Li Lithium	9 $_{4}$Be Beryllium											10.8 $_{5}$B Boron	12 $_{6}$C Carbon	14 $_{7}$N Nitrogen	16 $_{8}$O Oxygen	19 $_{9}$F Fluorine	20.2 $_{10}$Ne Neon
Period 3	23 $_{11}$Na Sodium	24.3 $_{12}$Mg Magnesium											27 $_{13}$Al Aluminium	28.4 $_{14}$Si Silicon	31 $_{15}$P Phosphorus	32.1 $_{16}$S Sulphur	35.5 $_{17}$Cl Chlorine	39.9 $_{18}$Ar Argon
Period 4	39.1 $_{19}$K Potassium	40.1 $_{20}$Ca Calcium	45 $_{21}$Sc Scandium	47.9 $_{22}$Ti Titanium	50.9 $_{23}$V Vanadium	52 $_{24}$Cr Chromium	54.9 $_{25}$Mn Manganese	55.9 $_{26}$Fe Iron	58.9 $_{27}$Co Cobalt	58.7 $_{28}$Ni Nickel	63.5 $_{29}$Cu Copper	65.4 $_{30}$Zn Zinc	69.7 $_{31}$Ga Gallium	72.6 $_{32}$Ge Germanium	74.9 $_{33}$As Arsenic	79 $_{34}$Se Selenium	79.9 $_{35}$Br Bromine	83.8 $_{36}$Kr Krypton
Period 5	85.5 $_{37}$Rb Rubidium	87.6 $_{38}$Sr Strontium	88.9 $_{39}$Y Yttrium	91.2 $_{40}$Zr Zirconium	92.9 $_{41}$Nb Niobium	95.9 $_{42}$Mo Molybdenum	98.0 $_{43}$Te Technetium	101.1 $_{44}$Ru Ruthenium	102.9 $_{45}$Rh Rhodium	106.1 $_{46}$Pd Palladium	107.6 $_{47}$Ag Silver	112.4 $_{48}$Cd Cadmium	115 $_{49}$In Indium	119 $_{50}$Sn Tin	122 $_{51}$Sb Antimony	128 $_{52}$Te Tellurium	127 $_{53}$I Iodine	131 $_{54}$Xe Xenon
Period 6	133 $_{55}$Cs Caesium	137.3 $_{56}$Ba Barium	57-71 Lanthanide series	178.5 $_{72}$Hf Hafnium	180.9 $_{73}$Ta Tantalum	183.9 $_{74}$W Tungsten	186.2 $_{75}$Re Rhenium	190.2 $_{76}$Os Osmium	192.2 $_{77}$Ir Iridium	195.1 $_{78}$Pt Platnum	197.0 $_{79}$Au Gold	200.6 $_{80}$Hg Mercury	203 $_{81}$Tl Thallium	207 $_{82}$Pb Lead	209 $_{83}$Bi Bismuth	210 $_{84}$Po Polonium	210 $_{85}$At Astatine	222 $_{86}$Rn Radon
Period 7	223 $_{87}$Fr Francium	226 $_{88}$Ra Radium	89 - 103 Actinide series	267 $_{104}$Rf Rutherfordium	262 $_{105}$Db Dubnium	271 $_{106}$Sg Seaborgium	270 $_{107}$Bh Bohrium	277 $_{108}$Hs Hassium	278 $_{109}$Mt Meitherium	281 $_{110}$Ds Francium	282 $_{111}$Rg Francium	285 $_{112}$Cn Francium	286 $_{113}$Nh Nihonium	289 $_{114}$Fl Flerovium	289 $_{115}$Mc Moscovium	293 $_{116}$Lv Livermorium	294 $_{117}$Ts Tennessine	294 $_{118}$Og Oganesson

Lanthanide series	138.9 $_{57}$La Lanthanum	140.1 $_{58}$Ce Cerium	140.9 $_{59}$Pr Praseodymium	144.2 $_{60}$Nd Neodymium	145 $_{61}$Pm Promethium	150.4 $_{62}$Sm Samarium	152 $_{63}$Eu Europium	157.3 $_{64}$Gd Gadolium	158.9 $_{65}$Tb Terbium	162.5 $_{66}$Dy Dysprosium	164.9 $_{67}$Ho Holmium	167.3 $_{68}$Er Erbium	168.9 $_{69}$Tm Thulium	173 $_{70}$Yb Ytterbium	175 $_{71}$Lu Lutetium
Actinide series	227 $_{89}$Ac Actinium	232 $_{90}$Th Thorium	231 $_{91}$Pa Protactinium	238 $_{92}$U Uranium	237 $_{93}$Np Neptunium	244 $_{94}$Pu Plutonium	243 $_{95}$Am Americium	247 $_{96}$Cm Curium	247 $_{97}$Bk Berkelium	251 $_{98}$Cf Californium	252 $_{99}$Es Einsteinium	257 $_{100}$Fm Fermium	258 $_{101}$Md Mendelevium	259 $_{102}$No Nobelium	262 $_{103}$Lr Lawrencium

Glossary

Aftershock	a ground tremor caused by the repositioning of rocks after an earthquake
Amplifier	an electronic device used to increase the strength (power) of a signal fed into it
Amplitude	(of a wave) the maximum displacement of medium particle measured from its equilibrium position
Audibility range	the range of frequencies of sound waves that can be heard by a listener. A listener can be a human or other animals
Beat	an interference between two sound waves of slightly different frequencies, perceived as periodic variations in volume whose rate is the difference between the two frequencies
Binding energy	the energy required to hold the nucleons forming the nucleus.
Cloud chamber	a chamber used for detecting ionizing radiation by using water droplets
Conductor	a material through which electrons will readily flow, thus allowing the material to transfer electrical energy
Constellation	a group of stars which form a definite pattern in the sky
Diode	a semiconductor device made up of a single junction between an *n*-type and a *p*-type area. Diodes allow conduction of electric current in only one direction
Dopant	a substance, such as boron, added as impurity to a pure semiconductor material for the purpose of altering its conductive properties
Doping	the deliberate addition of a very small amount of impurities to a pure semiconductor crystal. The added impurities give the semiconductor an excess of conducting electrons or an excess of conducting holes

Echo	reflected sound wave that arrives with sufficient magnitude and time delay such that it is distinguishable from a sound wave received directly from a source
Eddy currents	current loops induced by a changing magnetic field that circulates in a body of a conductor
Electromagnetic induction	the production of a voltage in a conductor due to a change in the number of magnetic lines of force (flux linkages) passing through the conductor
Electromagnetic spectrum	the entire distribution of electromagnetic radiation according to the increasing or decreasing magnitude of frequency or wavelength
Electromagnetic wave	a wave that consists of an electric field and a magnetic field oscillating with the same frequency, but perpendicular to each other
Electromotive force	the work or energy that causes a flow of an electric current through a circuit. It is also the potential difference (voltage) between two electrodes of a cell
Electron gun	an electrical component that produces an electron beam that has a precise kinetic energy and is most often used in televisions and monitors which use cathode-ray tube technology
Epicentre	the point on the Earth's surface directly above the focus (hypocentre) of an earthquake
Fission	the splitting of unstable nuclei to form smaller daughter nuclei, and release energy
Flux	a measure of how much of a field passes through a perpendicular unit surface area
Forward bias	arrangement of a semi-conductor circuit such that current flows when the circuit is completed
Frequency	(of a wave) the number of complete oscillations (cycles) that occur per unit of time
Fundamental frequency	the lowest frequency produced by a particular musical instrument.
Fusion	joining of small atomic nuclei together to create a larger nucleus, releasing enormous amount of energy

Galaxy	a gravitationally bound system of stars, usually numbering billions, that orbit a common mass and travel through the universe as a single unit
Geiger-muller counter	instrument for detecting ionising radiation emitted by radioactive nuclei
Generator	a device that converts mechanical energy into electrical energy by spinning conductors, usually copper wires, within a magnetic field
Global warming	the gradual increase in global temperatures as a result of greenhouse gases trapping longwave radiation emitted from the Earth's surface
Hard X-rays	X-rays with short-wavelength, with a very high penetrating power
Harmonic	a whole-number multiple of a fundamental frequency
Hypocentre	the point within the Earth from where an earthquake occurs and the accompanying seismic waves originate
Infrasonic	(of sound) having a frequency lower than the range of frequencies perceptible to the human ear, usually below 20 Hz
Insulator	a material that does not transmit sound, heat (thermal insulator), or electricity (electrical insulator)
Interference	the addition (superposition) of two or more waves that results in a new wave pattern
Lava	molten rock (magma) that reaches the Earth's surface during a volcanic eruption
Light-year	the distance that light travels in one year
Line of sight	a characteristic of electromagnetic radiation in which two stations can only transmit and receive data signals when they are in direct view of each other with no obstacle in between
Longitudinal wave	a wave in which the particles of the medium vibrate in a direction parallel to the direction of propagation of the wave
Loudness	a measure of how audible a sound is, depending on the frequency of sound and the sensitivity of one's ear. Loudness differs from one person to another

Magma	molten rock containing liquids, crystals and dissolved gases that forms within the upper part of the Earth's mantle
Magnetic flux	the total number of magnetic field lines of force produced by a magnet passing through a unit area perpendicularly
Maria	the dark spots on the Moon, which are covered by a type of volcanic rock called basalt
Mechanical waves	a group of waves that require a material medium for them to transfer energy
Meteoroids	objects found in space that range in size from dust grains to size close to small asteroids
Moon	a natural object that is in orbit around a planet
Mutual induction	the production of an electromotive force in a circuit resulting from a change of current in a neighboring circuit
Nocturnal	the condition of being active during the night and not during the day
Node	a point on a stationary wave with zero amplitude
Overtone	all of the frequencies present in a tone besides the fundamental frequency
Period	the time it takes for a wave to make one complete oscillation (cycle)
Phosphors	solid materials that emit light when exposed to radiation such as electron beam or ultra violet light
Pitch	the frequency of a sound as perceived by the human ear. A high-pitch sound corresponds to a high-frequency sound wave while a low-pitch sound corresponds to a low-frequency sound wave
Plasma	an electrically conducting medium in which there are roughly equal numbers of positively and negatively charged particles, produced when the atoms in a gas become ionized
Planet	a celestial body orbiting a star that is massive enough for gravity to have squeezed it into spherical shape and has cleared its neighbourhood
Radiation dose	a measure of the amount of exposure to nuclear radiation

Radioisotope	an unstable isotope of an element that decays or disintegrates spontaneously, emitting radiation
Rectification	the process of converting alternating current to a direct current
Resonance	the vibration of an object at its natural frequency on being set into vibration by the vibrations of another adjoining or interconnected object
Reverberation	accumulation of sound waves in a space caused by numerous reflections of sound or signal
Reverse bias	arrangement of a semi conductive circuit such that no current flows when the circuit is completed
Rock	naturally occuring and coherent aggregate of one as more minerals
Seismic waves	a series of progressive disturbances that are propagated through the Earth transmitting the energy released from an earthquake
Solar mass ($M_\odot$)	a standard unit of mass in astronomy, equal to approximately 2×10^{30} kg
Solar system	a system consisting of a star and other celestial objects revolving around it. Our solar system consists of the Sun and the planets, their satellites, the minor planets, comets, meteoroids and other objects revolving around the Sun.
Solar wind	a stream of charged particles released from the upper part of the Sun, called the corona
Solenoid	a coil of insulated wire which is used to produce a concentrated and uniform magnetic field when connected in an electrical circuit
Spark counter	a device used for detecting and counting ionisation of air caused by alpha radiation
Stationary wave	a wave that remains on a constant position, with fixed nodes and antinodes
Star	a celestial body made up of hot gases that radiates energy derived from thermonuclear reactions in the interior
Tectonic plate	a large, distinct, consolidated fragment of the Earth's crust and the upper mantle that moves about over the mantle, causing mountains to grow where plates collide and allowing new rock to form as hot magma rises to the surface where plates move apart

Temperature	the degree of hotness or coldness of a body. It is a measure of the average kinetic energy of the individual particles (atoms or molecules) that make up a body
Tide	the periodic rise and fall of the level of ocean waters which results from gravitational attraction of the Moon and Sun acting upon the rotating Earth
Timbre	the texture or characteristic that defines a sound produced by a specific instrument. It is also called quality
Transducer	a device that converts one type of energy to another, such as electrical energy to magnetic energy and vice versa
Transistor	a solid state electronic device that is used to control flow of electricity in electronic equipment and usually consists of a small block of a semiconductor (such as germanium) with at least three electrodes
Transverse wave	a wave in which the particles of the medium vibrate in a direction that is perpendicular to the direction of propagation of the wave
Tsunami	one or a series of huge sea waves caused by earthquakes or other large-scale disturbances of the ocean floor
Ultrasonic	(of sound) having a frequency higher than the range of frequencies perceptible to the human ear, usually 20 kHz or higher
Valence band	the highest range of electron energies in which electrons in an atom are normally present at absolute zero temperature
Valence electrons	are the electrons in the outermost shell of an atom which are responsible for chemical reaction
Volcano	an opening in the Earth's crust through which molten magma and gases force their way onto the Earth's surface. It also refers to the solid structure created from the lava, gases and hot particles that accompany a volcanic eruption
Wave	a periodic disturbance that propagates through space (medium) and with time, usually transferring energy from one point to another
Zodiac	a region of the sky in which the Sun, Moon and planets appear to travel along their paths

Bibliography

Abbott, A.F. (1989). *Physics (5th ed.)*. Oxford: Heinemann Educational.

Breithaupt, J. (2013). *Physics in Context for Cambridge International A Level*. United Kingdom: Nelson Thornes.

Bond, P. (2020) *Exploring the Solar System* United Kingdom, Wiley Blackwell

Chaisson, E., & McMillan, S. (2014). *Astronomy Today (8th ed.)*. Sansome St, San Fransico: Pearson Education.

Duncan, T., & Kennett, H. (2014). *Cambridge IGCSE Physics (3rd ed.)*. London, England: Hodder Education.

Gibilisco, S. (2003). *Astronomy Demystified*. New York, Chicago: McGraw-Hill Companies, Inc.

Nelkon, M. (1990). *Principles' Physics (8th ed.)*. London, United Kingdom: Peason Education Ltd.

Lang, K.R. (2013) *Essential Astrophysics*. New York, Springer Heidelberg.

South Carolina State University & Ministry of Education and Vocational Training, Zanzibar. (2008). *Physics for Secondary Schools Form 4 (Revised edition)*. Dar es Salaam: Oxford University Press Tanzania Ltd.

Tanzania Institute of Education. (1995). *Physics for Secondary Schools Book Four*. Dar es Salaam: Publishers and Distributors Ltd.

Taylor, M. (1999). *Macmillan Secondary Physics*. London and Oxford: Macmillan education Ltd.

Tanzania Institute of Education. (2019). *Physics for Secondary Schools Book Five*. Dar es Salaam: Tanzania Institute of Education.

Index

A

alpha particles 87, 88, 89, 96, 97, 98, 108, 112
alpha particles, 96, 112
amplitude 2, 4, 6, 7, 8, 9, 14, 15, 22, 23, 26, 38, 47, 48, 149, 150, 197, 220
Answer 214
antinode 28, 30, 34, 36
apogee 177, 178, 183
asthenosphere 189
astronomy 123, 157, 158, 183, 221
atmosphere 5, 45, 46, 172, 173, 193, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 221
atom 44, 86, 87, 88, 89, 90, 91, 92, 93, 112, 113, 115, 116, 120, 131, 133, 134, 135, 154, 222
atomic number 90, 91, 92, 93, 94, 107, 112
audibility range 21

B

beats 38
beta particles 94, 96, 97, 100, 112, 114
binding energy 93, 94, 108, 115

C

camera 45, 99, 100
chain reaction 109, 113
cloud chamber 46, 97, 99, 100
coil 55, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 74, 75, 76, 77, 78, 79, 80, 81, 83, 84, 85, 200, 221
constructive interference 15, 16, 48
corona 221
craters 178, 179, 194
crust 177, 187, 188, 189, 191, 192, 194, 196, 211, 212, 213, 221, 222
current 39, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 83, 84, 85, 99, 116, 117, 123, 124, 125, 127, 130, 131, 132, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 156, 212, 217, 218, 220, 221

D

deflection system 116
density 6, 8, 30, 49, 172, 185, 187, 189, 194, 195, 196
destructive interference 15, 16, 47, 48
diffraction of waves 18, 50

E

ear 12, 21, 22, 23, 26, 47, 219, 220, 222
earth 45, 195
echo 23, 24, 25, 49
effusive eruption 193
electromagnetic spectrum 40, 41, 43, 44, 48, 49
electromagnetic waves 5, 7, 39, 40, 45, 47, 48, 122, 202, 205
electromagnetism 51, 52, 63
electron 44, 45, 87, 90, 91, 94, 96, 116, 118, 119, 120, 121, 123, 124, 125, 132, 134, 135, 136, 139, 153, 154, 218, 222
evening star 170, 175, 185
exoplanet 171, 172, 175
expansion 19, 209
expansion, 19
eye 39, 122, 169, 170, 174

F

Faraday's Law 71
flux 61, 62, 63, 64, 65, 66, 67, 68, 69, 71, 74, 75, 78, 80, 218, 220
focal point 11
forced vibrations 33
frequency 3, 4, 7, 13, 16, 17, 19, 20, 21, 22, 23, 26, 28, 29, 30, 31, 32, 33, 35, 37, 38, 39, 40, 41,

46, 47, 48, 49, 50, 72, 116, 120, 143, 192, 193, 202, 211, 218, 219, 220, 221, 222
fundamental note 28, 29, 33, 36

G

gamma radiation 112
gamma rays 39, 40, 46, 48, 50, 93, 95, 97, 100, 122
generator 16, 17, 26, 71, 72, 73, 74, 81, 110
global warming 194, 206, 208, 209, 210, 211, 213
gravity 172, 176, 177, 220
greenhouse gases 110, 204, 206, 207, 208, 209, 210, 212, 219

H

half-life 101, 102, 103, 112, 113, 114
harmonic 28, 29, 30, 33, 34, 37, 38

I

incident wave 8, 9, 17, 18
index finger 58, 80

L

Lenz,s Law 71
longitudinal waves, 47

M

magnetic field, 52, 54, 57, 61, 62, 69, 71, 72, 73, 96
magnetic flux 67, 68, 69, 71, 80
mass number 90, 91, 92, 94, 107, 111, 112, 113
mechanical waves 5, 20, 39, 47
microwaves 39, 42, 43, 48
musical sounds 26, 27
mutual induction 68, 74

N

neutron 46, 90, 91, 92, 94, 108, 109, 110, 112
nuclear 46, 50, 73, 90, 91, 93, 94, 95, 96, 97, 98, 99, 104, 105, 107, 108, 109, 110, 111, 112, 113, 114
nuclide 91, 94, 104, 113

O

overtone 26, 28, 30, 34, 37, 39

P

perigee 183
period 3, 4, 47, 49, 103, 106, 161, 172, 173, 177, 181
photocells 44
photosynthesis 44, 206
Physics i
pitch 26, 48, 50, 220
primary coil 68, 69, 70, 74, 75, 76, 77, 79
principle of superposition 14, 38
proton 90, 91, 92, 93, 94

R

radiation 39, 40, 42, 43, 45, 46, 47, 48, 49, 50, 93, 95, 97, 98, 99, 100, 101, 105, 106, 107, 111, 112, 114, 120, 126, 204, 205, 206, 207, 209, 211, 218, 219, 221
radioactive decay 46, 95, 112, 114
radioactivity 86, 93, 104, 105, 113, 114
radio waves 5, 19, 39, 41, 42, 48, 50, 122
rarefactions 6
Refraction of waves 12
resonance 33, 35, 36, 37, 39, 49, 50, 126
reverberation 23, 24
ripple tank 10, 12, 15, 16, 17, 49

S

self-induction 68
shooting stars 173
solenoid 54, 55, 57, 65, 80, 82
sonometer 30, 50
sound 5, 6, 7, 12, 16, 18, 20, 21, 22, 23, 24, 25, 26, 30, 31, 33, 34, 35, 36, 37, 38, 39, 47, 49, 50,

146, 148, 149, 198, 199, 217, 218, 219, 220, 221, 222
spark counter 97, 103, 112
spectrum 40, 218
speed 3, 5, 7, 8, 12, 14, 19, 23, 24, 25, 35, 39, 40, 42, 47, 62, 97, 122, 173, 189, 196
stationary waves 30
strong force 93

T

temperature 7, 23, 24, 25, 35, 43, 44, 49, 115, 116, 117, 123, 128, 131, 132, 133, 148, 149, 153, 155, 172, 187, 188, 189, 201, 202, 203, 204, 205, 207, 208, 209, 211, 212, 213, 222
thermal imaging 43, 44
thermometer 35
transformer 74, 75, 76, 77, 78, 79, 83, 84, 85, 145
transmission 42, 74, 78
transmitted wave 8, 9
transverse waves 6, 7, 27, 39, 47

U

ultrasonic 21
ultraviolet 39, 45, 48, 205, 206

V

visible light 5, 39, 43, 44, 45, 48, 49, 50, 120
vitamin 45, 49
voltage 68, 69, 72, 74, 75, 76, 77, 78, 79, 83, 84, 97, 99, 117, 119, 120, 121, 124, 134, 136, 137, 138, 139, 141, 142, 143, 144, 145, 146, 148, 149, 150, 154, 218

W

water waves 5, 7, 10, 11, 12, 13, 16, 17, 18, 20, 49, 200
wavelength 2, 4, 6, 7, 8, 12, 13, 17, 18, 19, 20, 29, 34, 35, 36, 37, 38, 40, 41, 42, 49, 50, 141, 200, 207, 218
waves 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 30, 38, 39, 40, 41, 43, 44, 46, 47, 48, 49, 106, 195, 196, 197, 199, 200, 203, 205, 209, 211, 212, 217, 219, 220, 221, 222

Z

zodiac 163, 164, 165