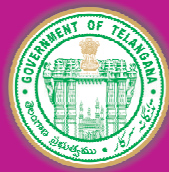


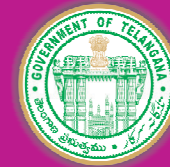
# INTERMEDIATE PHYSICS



Government of Telangana



[www.telanganaopenschool.org](http://www.telanganaopenschool.org)



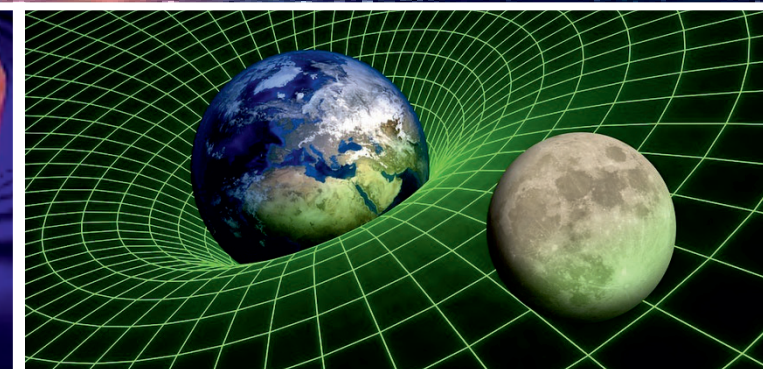
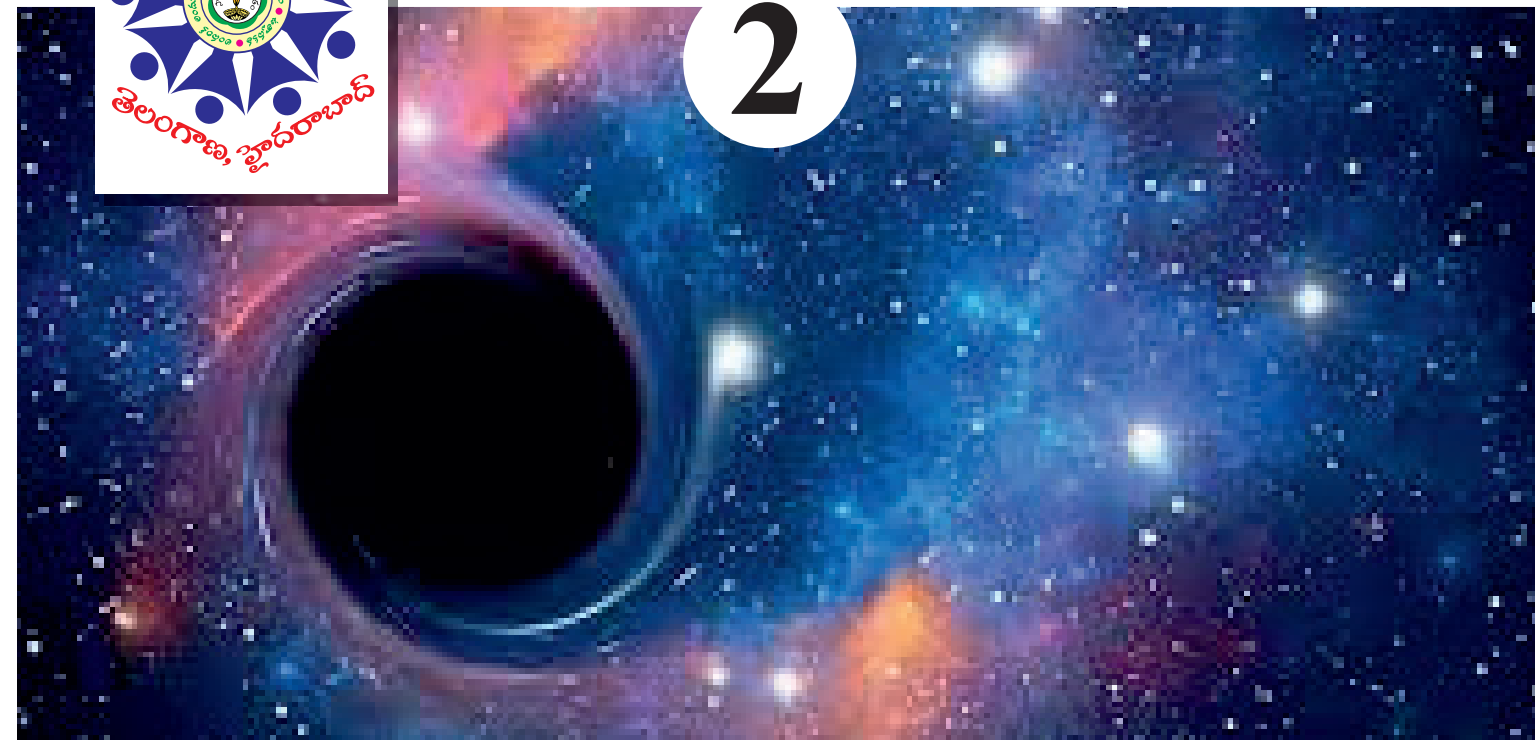
Government of Telangana

312

# INTERMEDIATE PHYSICS



2



TELANGANA OPEN SCHOOL SOCIETY, HYDERABAD

# 312

# PHYSICS

## 2

### Chief Advisor

**Smt. Vakati Karuna, IAS,**  
Secretary to Government, Education Department,  
Govt. of Telangana, Hyderabad.

### Editors

**Dr. B. Appa Rao**  
Professor Emeritus,  
JNTUH & Professor (Retd.),  
Department of Physics,  
Osmania University, Hyderabad.

**Dr. G. Bhikshamaiah**  
Retd. Professor of Physics,  
Department of Physics,  
Osmania University,  
Hyderabad.

### Textbook Printing Council

**Smt. A. Sridevasena, IAS,**  
Director, School Education, Telangana, Hyderabad.

**Sri P.V. Srihari, Director,**  
TOSS, Telangana, Hyderabad.

**Sri S. Srinivasa Chary**  
Director, Textbook Press, Telangana, Hyderabad.

### Chief Coordinator

**Sri M. Somi Reddy**  
Joint Director, TOSS, Telangana, Hyderabad.



Printed by

**Telangana Open School Society (TOSS), Hyderabad.**



© *Telangana Open School Society*

**First Published : 2023**

**All Rights Reserved**

No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means without the prior permission, in writing of the publisher, nor be otherwise circulated in any form of binding or cover.

---

*Printed in India*  
at the **Telangana Govt. Text Book Press,**  
Mint Compound, Khairathabad, Hyderabad, Telangana.

## FOREWORD

Providing education to children is a fundamental right and is essential for the overall development of society. The Government of Telangana plays a crucial role in ensuring that education is accessible to all, often establishing institutions like the Telangana Open School Society (TOSS) to serve children who may face challenges in accessing formal education due to various reasons.

In order to deliver quality education to students studying Intermediate Education in the Telangana Open School Society, starting from the 2023 academic year, we have undertaken the task of revising our textbooks to align them with the changing social landscape and to incorporate the fundamental principles of the National Education Policy 2020. The guidelines outlined in this policy aim to enhance the overall learning experience and cater to the diverse needs of our students. Unlike the earlier textbooks that primarily contained questions and answers, TOSS has taken a student-centric approach in designing these textbooks, taking into consideration the various learning styles and needs of our students. This approach encourages active engagement and participation in the learning process.

This Physics textbook is divided into two volumes, encompassing a total of 29 chapters. It has been thoughtfully crafted to foster the understanding and appreciation of this remarkable discipline. Each volume serves as a gateway to the intriguing and multifaceted realm of physics.

Volume-1 lays the essential foundation for students' journey into the world of physics. Mechanics, with its principles of motion, forces, and energy, provides the framework for comprehending how objects interact with one another. Heat and thermodynamics delve into the fascinating world of temperature, heat transfer, and the laws governing energy transformations.

Volume-2 will guide the learner through the enigmatic universe of light, exploring topics like reflection, refraction, and the formation of images, and delve deeper into the enigmas of the physical world. Electricity and magnetism will ignite the imagination of learners as they explore the principles of electric circuits, magnetic fields, and electromagnetic waves. Atoms and nuclei will unveil the intricacies of the atomic realm, from quantum mechanics to nuclear physics. Semiconductors and communication systems will introduce students to the technology that drives our modern world, from transistors to telecommunications.

We are indeed very grateful to the Government of Telangana and the Telangana State Board of Intermediate Education. Special thanks go to the editors, authors, co-coordinator, teachers, lecturers, and DTP operator who tirelessly contributed their services to create this textbook.

Date : 2023  
Place : Hyderabad.

Director,  
TOSS, Hyderabad.



## Textbook Development Committee

**Dr. Md. Shareefuddin**

Professor of Physics  
Department of Physics  
University College of Science  
Osmania University, Hyderabad.

**Dr. E. Rukmini**

Associate Professor  
Department of Physics & Electronics  
Sarojini Naidu Vanita Maha Vidyalaya  
Hyderabad.

**Dr. N. V. Prasad**

Associate Professor  
Department of Physics  
Osmania University, Hyderabad.

**Dr. S. Vedavyas**

Research Assistant (Physics)  
Telugu Akademi  
Himayatnagar, Hyderabad.

**Dr. Y. Vasudeva Reddy**

Assistant Professor  
Department of Physics  
Government City College (A)  
Nayapul, Hyderabad.

**Sri K. Subba Rao**

Associate Professor  
Department of Physics & Electronics  
Sarojini Naidu Vanita Maha Vidyalaya  
Hyderabad.

**Dr. J. Chinna Babu**

Assistant Professor  
Department of Physics  
Vivekananda Govt. Degree College  
Vidyanagar, Hyderabad.

**Mrs. B. Anuradha**

Associate Professor  
Department of Physics & Electronics  
Sarojini Naidu Vanita Maha Vidyalaya  
Hyderabad.

**Dr. N. Thirumal Reddy**

Assistant Professor of Physics  
Government Degree College  
Medak.

**Sri D. Anantha Ramakrishna**

Principal  
Govt. Junior College  
Dharmaram, Peddapalli.

### Coordinators

**Sri B. VenkateswaraRao**

State Coordinator, TOSS, Telangana,  
Hyderabad.

**Smt K. Madhavi**

State Coordinator, TOSS, Telangana,  
Hyderabad.

*Cover Design :* **Sri K. Sudhakar Chary**, S.G.T.,  
UPPS Mylaram, Rayaparathi (M), Warangal Rural.

*DTP, Page Layout & Design by :*  
**Smt. K. PAVANI**, Hyderabad.

## A Word With You

Dear Learner

We are delighted to welcome you to the world of open and distance learning through TOSS. Your decision to embark on this educational journey as an Open and Distance learner is commendable, and we are thrilled that you have chosen physics as one of your subjects of study.

With great pleasure and enthusiasm, we present to you this comprehensive physics textbook, meticulously designed for senior secondary course in open and distance learning. Physics is a subject that unveils the mysteries of the universe, offering profound insights into the fundamental laws governing the cosmos. Our textbook, spanning two volumes and encompassing 29 chapters, has been thoughtfully crafted to foster your understanding and appreciation of this remarkable discipline.

As you set out on this educational voyage, you will explore the captivating domains of mechanics, heat and thermodynamics, optics, electricity and magnetism, atoms and nuclei, semiconductors, and communication systems. Each volume serves as a gateway to the intriguing and multifaceted realm of physics.

Volume-1, consisting of 14 chapters, lays the essential foundation for your journey into the world of physics. Mechanics, with its principles of motion, forces, and energy, provides the framework for comprehending how objects interact with one another. Heat and thermodynamics delve into the fascinating world of temperature, heat transfer, and the laws governing energy transformations.

Volume-2, with 15 chapters, will lead you through the enigmatic universe of light, exploring topics like reflection, refraction, and the formation of images, delves deeper into the enigmas of the physical world. Electricity and magnetism will electrify your imagination as you explore the principles of electric circuits, magnetic fields, and electromagnetic waves. Atoms and nuclei will unveil the intricacies of the atomic realm, from quantum mechanics to nuclear physics. Semiconductors and communication systems will introduce you to the technology that drives our modern world, from transistors to telecommunications.

Our textbook is tailored with your unique learning journey in mind, recognizing the distinctive challenges and opportunities that open and distance learning offers. Clear explanations, illustrative diagrams, and practical examples have been incorporated to make physics accessible and engaging, even when you are learning independently. Each chapter acts as a stepping stone, building upon previously introduced concepts, ensuring a coherent and structured learning experience.

Physics transcends being a mere subject; it is a path to discovery. It empowers you to explore the universe, from the vast expanses of space to the tiniest particles that constitute matter. Physics nurtures critical thinking, problem-solving abilities, and a profound appreciation for the natural world's beauty.

As you progress through these pages, remember that physics is not just a collection of equations and theories; it is a tool that equips you to understand and shape the world around you. Embrace the challenges, ask questions, and never cease to wonder. Your journey through these volumes will not only equip you with the knowledge and skills to excel academically but also inspire a lifelong passion for the marvels of physics.

We extend our heartfelt best wishes as you embark on this educational odyssey. May this textbook be your trusted companion and guide on your quest for knowledge. If you encounter any difficulties or have suggestions, please do not hesitate to reach out to us.

Sincerely,

The Curriculum Design and Course Development Team

(312)  
**PHYSICS**

**VOLUME - 1**

1. Units and Measurements
2. Motion in a Straight Line
3. Motion in a Plane
4. Newton's Laws of Motion
5. Work, Energy and Power
6. System of Particles & Rotatory Motion
7. Simple Harmonic Motion
8. Gravitation
9. Mechanical Properties of Solids
10. Mechanical Properties of Fluids
11. Thermal Properties of Matter
12. Thermodynamics
13. Kinetic Theory of Gases
14. Sound Waves

\*\*\*\*\*



**(312)**  
**PHYSICS**

**VOLUME - 2**

15. Ray Optics and Optical Instruments
16. Wave Optics
17. Electrical Charges and Fields
18. Electrostatic Potential and Capacitance
19. Current Electricity
20. Magnetism
21. Moving Charges and Magnetism
22. Electromagnetic Induction
23. Alternating Currents
24. Electromagnetic Waves
25. Dual Nature of Radiation and Matter
26. Structure of the Atom
27. Nuclear Physics
28. Semiconductor Devices
29. Basics of Communication

\*\*\*\*\*

## CONTENTS

15.	Ray Optics and Optical Instruments	1 - 19
16.	Wave Optics	20 - 47
17.	Electrical Charges and Fields	48 - 74
18.	Electrostatic Potential and Capacitance	75 - 100
19.	Current Electricity	101 - 138
20.	Magnetism	139 - 152
21.	Moving Charges and Magnetism	153 - 176
22.	Electromagnetic Induction	177 - 198
23.	Alternating Currents	199 - 221
24.	Electromagnetic Waves	222 - 234
25.	Dual Nature of Radiation and Matter	235 - 251
26.	Structure of the Atom	252 - 264
27.	Nuclear Physics	265 - 296
28.	Semiconductor Devices	297 - 343
29.	Basics of Communication	344 - 363

\*\*\*\*\*



# RAY OPTICS AND OPTICAL INSTRUMENTS

## INTRODUCTION

We get most of the information about the surrounding world through our eyes. But as you know, an unaided eye has limitations; objects which are too far like stars, planets etc. appear so small that we are unable to see their actual details. Similarly, objects which are too small, e.g. pollen grains, bacteria, viruses etc., remain invisible to the unaided eyes. Moreover, our eyes do not keep a permanent record of what they see, except what is retained by our memory. You may therefore ask the question: How can we see very minute and very distant objects? The special devices meant for this purpose are called **optical instruments**.

In this lesson you will study about two important optical instruments, namely, a microscope and a telescope. As you must be knowing, a microscope magnifies small objects while a telescope is used to see distant objects. The design of these appliances depends on the requirement. (The knowledge of image formation by the mirrors and lenses, which you have acquired in previous classes, will help you understand the working of these optical instruments.) The utility of a microscope is determined by its magnifying power and resolving power. For a telescope, the key word is *resolving power*. You must have read about Hubble's space telescope, which is being used by scientists to get details of far off galaxies and search for a life-sustaining planet beyond our solar system.

## OBJECTIVES

After studying this lesson, you should be able to

- *explain the working principle of simple and compound microscopes;*
- *derive an expression for the magnifying power of a microscope;*
- *distinguish between linear and angular magnifications;*
- *explain the working principle of refracting and reflecting telescopes;*
- *calculate the resolving power of an eye, a telescope and a microscope.*

## 15.1) MICROSCOPE

In previous classes you have learnt about image formation by mirrors and lenses. If you take a convex lens and hold it above this page, you will see images of the alphabets/ words. If you move the lens and bring it closer and closer to the page, the alphabets printed on it will start looking enlarged. This is because their enlarged, virtual and erect image is being formed by the lens. That is, it is essentially acting as a magnifying glass or simple

microscope. You may have seen a doctor, examining measles on the body of a child or a dentist using it to examine a cavity in a tooth. Watch makers and jewellers use it to magnify small components of watches and fine jewellery work. You can take a convex lens and focus the sunlight on a small piece of paper. You will see that after some time, the piece of paper start burning. A convex lens can, therefore start a fire. That is why it is dangerous to leave empty glass bottles in the woods. The sunlight falling on the glass bottles may get focused on dry leaves in the woods and set them on fire. Sometimes, these result in wildfires, which destroy large parts of a forest and/or habitation. Such fires are quite common in Australia, Indonesia and U.S.

As a simple microscope, a convex lens is satisfactory for magnifying small nearby objects upto about twenty times their original size. For large magnification, a compound microscope is used, which is a combination of basically two lenses. In a physics laboratory, a magnifying glass is used to read vernier scales attached to a travelling microscope and a spectrometer.

While studying simple and compound microscopes, we come across scientific terms like (i) near point, (ii) least distance of distinct vision, (iii) angular magnification or magnifying power, (iv) normal adjustment etc. Let us first define these.

- (i) **Near point** is the distance from the eye for which the image of an object placed there is formed (by eye lens) on the retina. The near point varies from person to person and with the age of an individual. At a young age (say below 10 years), the near point may be as close as 7-8 cm. In the old age, the near point shifts to larger values, say 100-200 cm, or even more. That is why young children tend to keep their books so close where as the aged persons keep a book or newspaper far away from the eye.
- (ii) **Least distance of distinct vision** is the distance upto which the human eye can see the object clearly without any strain on it. For a normal human eye, this distance is generally taken to be 25 cm.
- (iii) **Angular magnification** is the ratio of the angle subtended by the image at the eye to the angle subtended by the object at the unaided eye when the object is placed at the least distance of distinct vision. It is also called the magnifying power of the microscope.
- (iv) **Normal Adjustment:** When the image is formed at infinity, least strain is exerted on the eye for getting it focused on the retina. This is known as normal adjustment.
- (v) **Linear magnification** is the ratio of the size of the image to the size of the object.
- (vi) **Visual angle** is the angle subtended by the object at human eye.

### 15.1.1 A Simple Microscope

A convex lens of short focal length which is used to see magnified image of a small object is called a simple microscope.

We know that when an object is placed between the optic center and the focus of a convex lens, its image is virtual, erect, and magnified and on the same side as the object. In practice, such a lens is held close to eye and the distance of the object is adjusted till

a clear image is formed at the least distance of distinct vision. This is illustrated in Fig. 15.1 (b), which shows an object  $AB$  placed between  $F$  and  $O$ . Its virtual image  $A'B'$  is formed on the same side as the object. The position of the object is so adjusted that the image is formed at the least distance of distinct vision ( $D$ ).

### Magnifying power of a simple microscope

Magnifying power of an optical instrument is the ratio of the angle subtended by the image at the eye to the angle subtended by the object seen directly, when both lie at the least distance of distinct vision or the near point. It is also called angular magnification and is denoted by  $M$ . Referring to Fig. 15.1 (a) and (b), the angular magnification of simple microscope is given by  $M = \frac{\angle A'OB'}{\angle AOB''} = \frac{\beta}{\alpha}$ . In practice, the angles  $\alpha$  and  $\beta$  are small.

Therefore, you can replace these by their tangents, i.e. write

$$M = \frac{\tan \beta}{\tan \alpha} \quad (15.1)$$

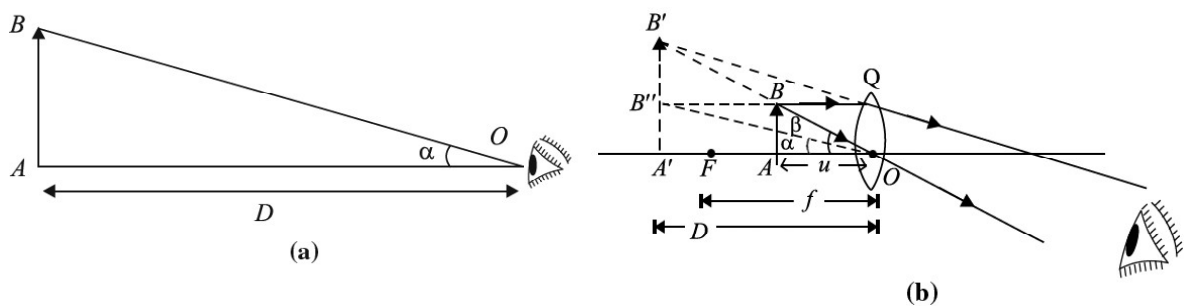


Fig. 15.1 : Angular magnification of a magnifying glass

From  $\Delta s A'OB'$  and  $AOB$ , we can write  $\tan \beta = \frac{A'B'}{AO} = \frac{A'B'}{D}$  and  $\tan \alpha = \frac{A'B''}{A'O} = \frac{AB}{D}$ .

On putting these values of  $\tan \beta$  and  $\tan \alpha$  in Eqn.(15.1), we get

$$M = \frac{A'B'}{D} \bigg/ \frac{AB}{D} = \frac{A'B'}{AB}$$

Since  $\Delta s AOB$  and  $A'OB'$  in Fig. 15.1(b) are similar, we can write

$$\frac{A'B'}{AB} = \frac{A'O}{AO} \quad (15.2)$$

Following the standard sign convention, we note that

$$A'O = -D$$

and

$$AO = -u$$

Hence, from Eqn. (15.2), we obtain

$$\frac{A'B'}{AB} = \frac{D}{u} \quad (15.3)$$

If  $f$  is the focal length of the lens acting as a simple microscope, then using the lens  $\left(\frac{1}{v} - \frac{1}{u} = \frac{1}{f}\right)$  formula and noting that  $v = -D, u = -u$  and  $f = f$ , we get

$$\frac{1}{-D} - \frac{1}{-u} = \frac{1}{f}$$

or 
$$-\frac{1}{D} + \frac{1}{u} = \frac{1}{f}$$

Multiplying both the sides by  $D$ , and rearranging terms, you can write

$$\frac{D}{u} = 1 + \frac{D}{f} \quad (15.4)$$

On combining Eqns. (15.3) and (15.4), we get

$$\frac{A'B'}{AB} = 1 + \frac{D}{f}$$

or 
$$M = 1 + \frac{D}{f} \quad (15.5)$$

From this result we note that lesser the focal length of the convex lens, greater is the value of the angular magnification or magnifying power of the simple microscope.

### Normal Adjustment

In normal adjustment, the image is formed at infinity. The magnifying power of the microscope is defined as the ratio of the angle subtended by the image at the eye to the angle subtended by the object at the unaided eye when the object is placed at  $D$ . Fig. 15.2.(a) shows that the object is placed at the least distance of distinct vision  $D$ .

The angles subtended by the object and the image at the unaided eye are  $\alpha$  and  $\beta$ , respectively. The magnifying power is defined as

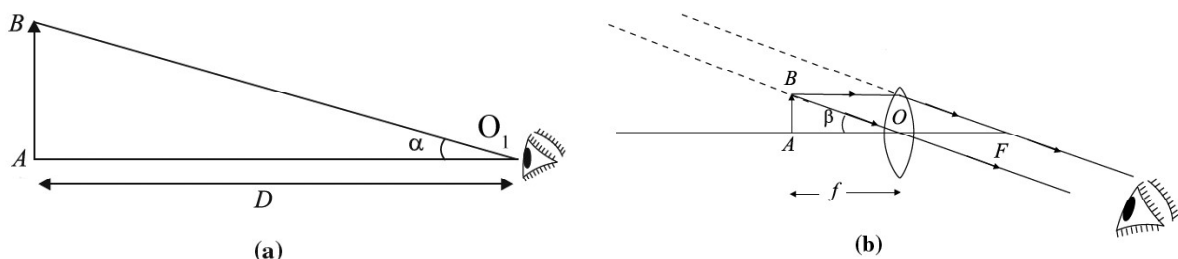


Fig. 15.2 : Image formation for normal adjustment

$$M = \frac{\beta}{\alpha}$$

In practice, the angles  $\alpha$  and  $\beta$  are small, and, as before, replacing these by their tangents, we get

$$\begin{aligned} M &= \frac{\tan \beta}{\tan \alpha} \\ \text{i.e.} \quad &= \frac{AB}{AO} \bigg/ \frac{AB}{AO_1} \\ &= \frac{AO_1}{AO} = \frac{D}{f} \end{aligned}$$

$$\text{or} \quad M = \frac{D}{f} \quad (15.6)$$

You may note that in the normal adjustment, the viewing of the image is more comfortable. To help you fix your ideas, we now give a solved example. Read it carefully.

### Example 15.1

Calculate the magnifying power of a simple microscope having a focal length of 2.5 cm.

#### Solution :

For a simple microscope, the magnifying power is given by Eqn.(15.5):

$$M = 1 + \frac{D}{f}$$

Putting  $D = 25$  cm and  $f = 2.5$ cm, we get

$$M = 1 + \frac{25}{2.5} = 1 + 10 = 11$$

### 15.1.2 A Compound Microscope

A compound microscope consists of two convex lenses. A lens of short aperture and short focal length faces the object and is called the **objective**. Another lens of short focal length but large aperture facing the eye is called the **eye piece**. The objective and eye piece are placed coaxially at the two ends of a tube.

When the object is placed between  $F$  and  $2F$  of the objective, it's a real, inverted and magnified image is formed beyond  $2F$  on the other side of the objective. This image acts as an object for the eye lens, which then acts as a simple microscope. The eye lens is so adjusted that the image lies between its focus and the optical center so as to form a magnified image at the least distance of distinct vision from the eye lens.



### Magnifying Power of a compound microscope

Magnifying power of a compound microscope is defined as the ratio of the angle subtended by the final image at the eye to the angle subtended by the object at unaided eye, when both are placed at the least distance of distinct vision. It is denoted by  $M$ . By referring to Fig. 15.3, we can write

$$M = \frac{\beta}{\alpha}$$

Since the angles  $\alpha$  and  $\beta$  are small, these can be replaced by their tangents, so that

$$M = \frac{\tan \beta}{\tan \alpha}$$

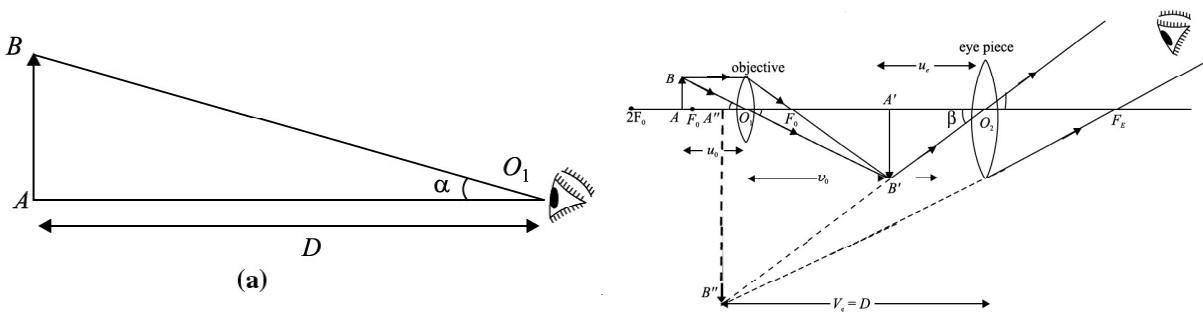


Fig. 15.3 : Image formation by a compound microscope when the final image is formed at the least distance of distinct vision.

$$M = \frac{A''B''}{D} \bigg/ \frac{AB}{D}$$

$$\Rightarrow M = \frac{A''B''}{AB} = \frac{A''B''}{A'B'} \cdot \frac{A'B'}{AB}$$

From similar  $\Delta s A''B''O_2$  and  $A'B'O_2$ , we can write

$$\frac{A''B''}{A'B'} = \frac{A''O_2}{A'O_2} = \frac{D}{u_e}$$

Also from similar  $\Delta s A'B'O_1$  and  $ABO$ , we have

$$\frac{A'B'}{AB} = \frac{\nu_o}{u_o}$$

Note that  $m_e = \frac{A''B''}{A'B'}$  defines magnification produced by eye lens and  $m_o = \frac{A'B'}{AB}$  denotes magnification produced by the objective lens. Hence

$$M = \frac{D}{u_e} \cdot \frac{v_o}{u_o} = m_e \times m_o \quad (15.7)$$

Recall the lens formula for eye lens, we can write

$$\frac{1}{v_e} - \frac{1}{u_e} = \frac{1}{f_e}$$

Multiply on both sides by  $v_e$  to get

$$\frac{v_e}{v_e} - \frac{v_e}{u_e} = \frac{v_e}{f_e}$$

$$\Rightarrow \frac{v_e}{u_e} = 1 - \frac{v_e}{f_e}$$

Since  $f_e$  is positive and  $v_e = -D$  as per sign convention, we can write

$$m_e = \frac{v_e}{u_e} = 1 + \frac{D}{f_e} \quad (15.8)$$

On combining Eqns. (15.7) and (15.8), we get

$$M = \frac{v_o}{u_o} \times \left( 1 + \frac{D}{f_e} \right)$$

In practice, the focal length of an objective of a microscope is very small and object  $AB$  is placed just outside the focus of objective. That is

$$v_o \approx f_o$$

Since the focal length of the eye lens is also small, the distance of the image  $A'B'$  from the object lens is nearly equal to the length of the microscope tube i.e.

$$u_e \approx L$$

Hence, the relation for the magnifying power in terms of parameters related to the microscope may be written as

$$M = \frac{L}{f_o} \left( 1 + \frac{D}{f_e} \right) \quad (15.10)$$

### Magnifying power in normal adjustment

In this case the image is formed at infinity. As discussed earlier, the magnifying power of the compound microscope may be written as

$$M = m_o \times m_e$$

$$= \frac{v_o}{u_o} \left( \frac{D}{f_e} \right)$$

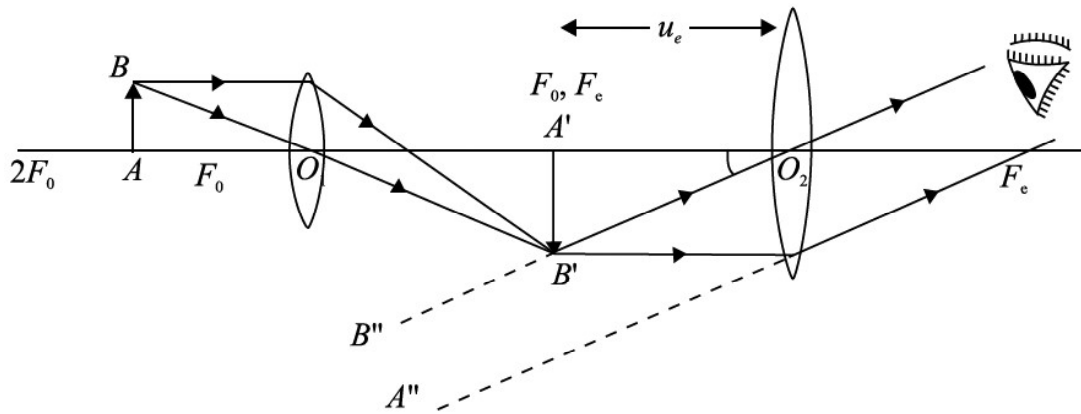


Fig. 15.4 : Compound microscope in normal adjustment

You may now like to go through a numerical example.

### Example 15.2

A microscope has an objective of focal length 2 cm, an eye piece of focal length 5 cm and the distance between the centers of two lenses is 20 cm. If the image is formed 30 cm away from the eye piece, find the magnification of the microscope.

### Solution :

For the objective,  $f_o = 2$  cm and  $f_e = 5$  cm. For the eye piece,  $v_e = -30$  cm, and  $f_e = 5$  cm. We can calculate  $u_e$  using the relation

$$\frac{1}{v_e} - \frac{1}{u_e} = \frac{1}{f_e}$$

On solving, you will easily obtain  $u_e = -\frac{30}{7}$  cm. For the objective lens

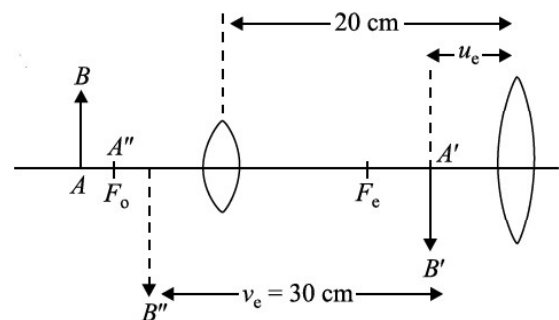
$$v_o = 20 - \frac{30}{7} = \frac{110}{7} \text{ cm}$$

Using the formula

$$\frac{1}{v_o} - \frac{1}{u_o} = \frac{1}{f_o}$$

We have

$$\frac{1}{110/7} - \frac{1}{u_o} = \frac{1}{2}$$



or 
$$u_o = \frac{110}{48} \text{ cm}$$

The magnifying power of the objective

$$m_o = \frac{v_o}{u_o} = \frac{\frac{110}{7}}{-\frac{110}{48}} = -\frac{48}{7}$$

The magnification due to the eye piece is

$$m_e = \frac{v_e}{u_e} = \frac{\frac{-30}{1}}{\frac{-30}{7}} = 7$$

Therefore, the magnification of the microscope is given by  $M = (m_o)(m_e)$

$$= \left(-\frac{48}{7}\right)(7) = -48$$

### Intext Questions 15.1

1. What is the nature of images formed by a (i) simple microscope (ii) Compound microscope?
2. Differentiate between the magnifying power and magnification?
3. The magnifying power of a simple microscope is 11. What is its focal length?
4. Suppose you have two lenses of focal lengths 100 cm and 4 cm respectively. Which one would you choose as the eyepiece of your compound microscope and why?
5. Why should both the objective and the eyepiece of a compound microscope have short focal lengths?

### 15.2) TELESCOPE

Telescopes are used to see distant objects such as celestial and terrestrial bodies. Some of these objects may not be visible to the unaided eye. The visual angle subtended by the distant objects at the eye is so small that the object cannot be perceived. The use of a telescope increases the visual angle and brings the image nearer to the eye. Mainly two types of telescopes are in common use : refracting telescope and reflecting telescope. We now discuss these.

### 15.2.1 Refracting Telescope

The refracting telescopes are also of two types:

- **Astronomical telescopes** are used to observe heavenly or astronomical bodies.
- **Terrestrial telescopes** are used to see distant objects on the earth. So it is necessary to see an erect image. Even Galilean telescope is used to see objects distinctly on the surface of earth.

An **astronomical telescope** produces a virtual and erect image. As heavenly bodies are round, the inverted image does not affect the observation. This telescope consists of a two lens system. The lens facing the object has a large aperture and large focal length ( $f_o$ ). It is called the *objective*. The other lens, which is towards the eye, is called the *eye lens*. It has a small aperture and short focal length ( $f_e$ ). The objective and eye-piece are mounted coaxially in two metallic tubes.

The objective forms a real and inverted image of the distant object in its focal plane. The position of the lens is so adjusted that the final image is formed at infinity. (This adjustment is called normal adjustment.) The position of the eye piece can also be adjusted so that the final image is formed at the least distance of distinct vision.

- (a) **When the final image is formed at infinity** (Normal adjustment), the paraxial rays coming from a heavenly object are parallel to each other and they make an angle  $\alpha$  with the principal axis. These rays after passing through the objective, form a real and inverted image in the focal plane of objective. In this case, the position of the eye piece is so adjusted that the final image is formed at infinity.

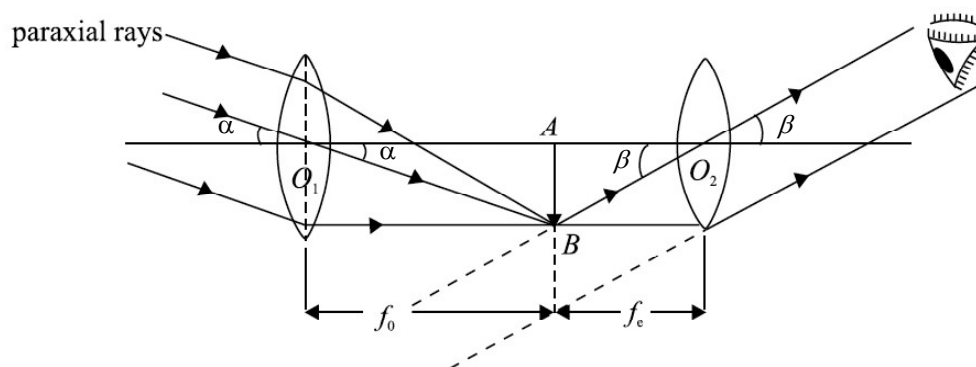


Fig. 15.6 : Working principle of an astronomical telescope

**Magnifying power** of a telescope is defined as the ratio of the angle subtended by the image at the eye as seen through the telescope to the angle subtended by the object at objective when both the object and the image lie at infinity. It is also called **angular magnification** and is denoted by  $M$ . By definition,

$$M = \frac{\beta}{\alpha}$$

Since  $\alpha$  and  $\beta$  are small, they can be replaced by their tangents. Therefore,

$$\begin{aligned}
 M &= \frac{\tan \beta}{\tan \alpha} \\
 &= \frac{AB / AO_2}{AB / AO_1} = \frac{AO_1}{AO_2} \\
 &= \frac{f_o}{f_e} \qquad (15.11)
 \end{aligned}$$

It follows that the magnifying power of a telescope in normal adjustment will be large if the objective is of large focal length and the eyepiece is of short focal length. The length of telescope in normal adjustment is  $(f_o + f_e)$ .

- (b) **When the final image is formed at the least distance of distinct vision**, the paraxial rays coming from a heavenly object make an angle  $\alpha$  with the principal axis. After passing through the objective, they meet on the other side of it and form a real and inverted image  $AB$ . The position of the eyepiece is so adjusted that it finally forms the image at the least distance of distinct vision.

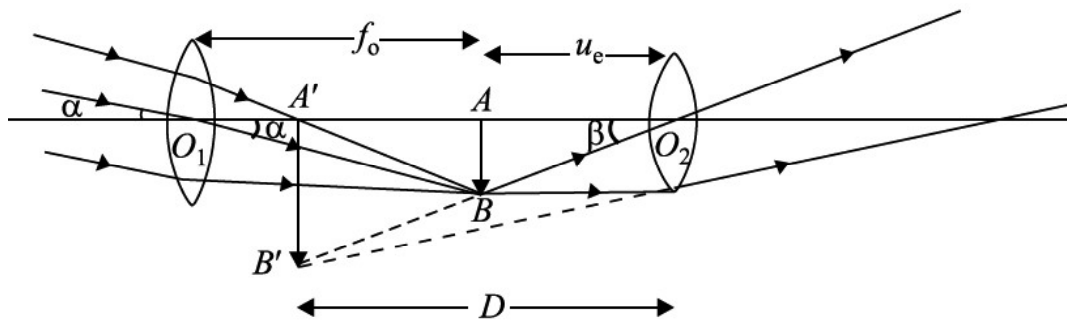


Fig. 15.7 : Image formed by a telescope at  $D$

**Magnifying power:** It is defined as the ratio of the angle subtended at the eye by the image formed at  $D$  to the angle subtended by the object lying at infinity:

$$\begin{aligned}
 M &= \frac{\beta}{\alpha} \\
 &\approx \frac{\tan \beta}{\tan \alpha} \\
 &= \frac{AB / AO_2}{AB / AO_1} = \frac{AO_1}{AO_2} \\
 &= \frac{f_o}{u_e} \qquad (15.12)
 \end{aligned}$$

Since  $\frac{1}{v_e} - \frac{1}{u_e} = \frac{1}{f_e}$  for the eyepiece, we can write

$$\frac{1}{u_e} = \frac{1}{v_e} - \frac{1}{f_e}$$

$$\approx \frac{-1}{f_e} \left( 1 - \frac{f_e}{v_e} \right)$$

or

$$M = \frac{f_o}{u_e} = \frac{-f_o}{f_e} \left( 1 - \frac{f_e}{v_e} \right) \quad (15.13)$$

Applying the new cartesian sign convention  $f_o = +f_o$ ,  $v_e = -D$ ,  $f_e = +f_e$ , we can write

$$M = \frac{f_o}{f_e} \left( 1 + \frac{f_e}{D} \right) \quad (15.14)$$

The negative sign of magnifying power of the telescope suggests that the final image is inverted and real. The above expression tells that the magnifying power of a telescope is larger when adjusted at the least distance of distinct vision to the telescope when focused for normal adjustment.

### 15.2.2 Reflecting telescope

A reflecting telescope is used to see distant stars and possesses large light-gathering power in order to obtain a bright image of even a faint star deep in space. The objective is made of a concave mirror, having large aperture and large focal length. This concave mirror, being parabolic in shape, is free from spherical aberration.

Before the reflected rays of light meet to form a real, inverted and diminished image of a distant star at the focal plane of concave mirror, they are intercepted and reflected by a plane Mirror  $M_1M_2$  inclined at an angle of  $45^\circ$  to the principal axis of the concave mirror. This plane mirror deviates the rays and the real image is formed in front of the eye piece, which is at right angle to the principal axis of concave mirror. The function of the eye-piece is to form a magnified, virtual image of the star enabling eye to see it distinctly.

If  $f_o$  is the focal length of the concave mirror and  $f_e$  is the focal length of eyepiece, the magnifying power of the reflecting telescope is given by

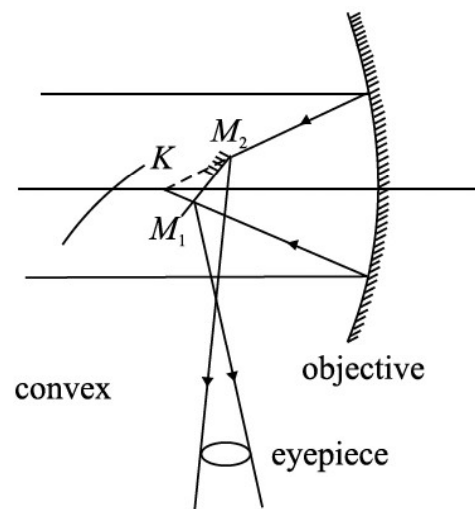


Fig. 15.8 : Newtonian Reflector

$$M = \frac{f_o}{f_e}$$

Further, if  $D$  is the diameter of the objective and  $d$  is the diameter of the pupil of the eye, the brightness ratio is given by

$$B = D^2/d^2$$

The other form of the reflecting telescope is shown in Fig. 15.9. It was designed by **Cassegrain**. In this case the objective has a small opening at its center. The rays from the distant star, after striking the concave mirror, are made to intercept at secondary mirror and the final image is viewed through the eye piece.

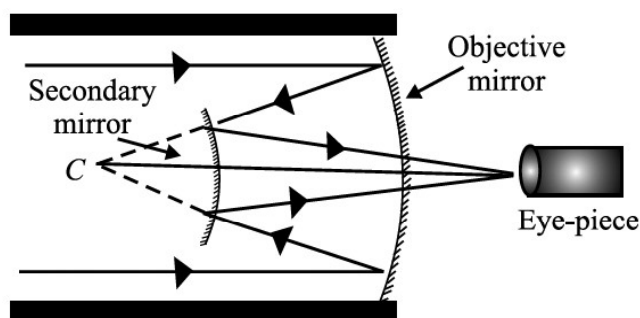


Fig. 15.9 : Cassegrain reflector

There are several advantages of a reflecting telescope over a refracting telescope.

- Since the objective is not a lens, the reflecting telescopes are free from chromatic aberration. Thus rays of different colours reaching the objective from distant stars are focussed at the same point.
- Since the spherical mirrors are parabolic mirrors, free from spherical aberration, they produce a very sharp and distinct image.
- Even a very faint star can be seen through the reflecting telescope because they have large aperture and have large light-gathering power. The brightness of the image is directly proportional to the area of the objective:

$$B \propto \frac{\pi D^2}{4}$$

where  $D$  is the diameter of the objective of the telescope. If  $d$  is the diameter of the pupil of the eye then brightness of the telescope  $B$  is defined as the ratio of light gathered by the telescope to that gathered by the unaided eye from the distant object

$$B = \frac{\pi D^2 / 4}{\pi d^2 / 4} = \frac{D^2}{d^2}$$

- In reflecting type of telescopes, there is negligible absorption of light.
- Large apertures of reflecting telescope enable us to see minute details of distant stars and explore deeper into space. That is why in recent years, astronomers have discovered new stars and stellar systems.



### Intext Questions 15.2

- How would the magnification of a telescope be affected by increasing the focal length of:
  - The objective
  - The eye piece
- If the focal length of the objective of a telescope is 50 cm and that of the eyepiece is 2 cm. What is the magnification?
- State one difference between a refracting telescope and reflecting telescope.
- What is normal adjustment?

### 15.3 RESOLVING POWER : THE RAYLEIGH'S CRITERION

In earlier lessons, you have seen that the image of a point source is not a point, but has a definite size and is surrounded by a diffraction pattern. Similarly, if there are two point sources very close to each other, the two diffraction patterns formed by the two sources may overlap and hence it may be difficult to distinguish them as separate by the unaided eye. The resolving power of an optical instrument is its ability to resolve (or separate) the images of two point objects lying close to each other. Rayleigh suggested that two images can be seen as distinct when the first minimum of the diffraction pattern due to one object falls on the central maximum of the other. This is called *Rayleigh's criterion*.

If we assume that the pupil of our eye is about 2 mm in diameter, two points can be seen distinctly separate, if they subtend an angle equal to about one minute of arc at the eye. **The reciprocal of this angle is known as the resolving power of the eye.**

Now let us calculate the resolving power of common optical instruments. We begin our discussion with a telescope.

#### 15.3.1 Resolving Power of a Telescope

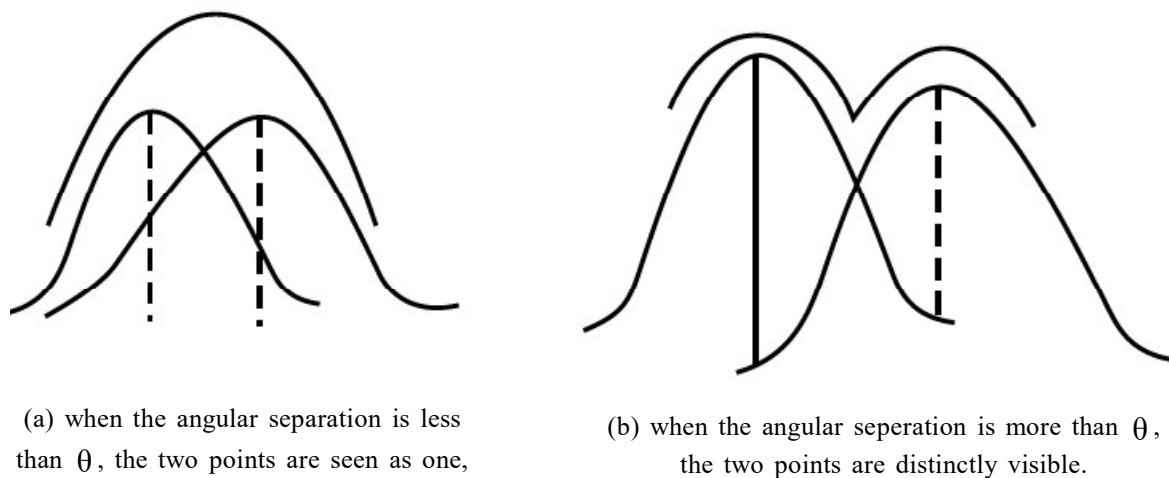


Fig. 15.10 : Rayleigh's criterion for resolution

The resolving power of a telescope is its ability to form separate images of two distant point objects situated close to each other. It is measured in terms of the angle subtended at its objective by two close but distinct objects whose images are just seen in the telescope as separate. This angle is called the **limit of resolution** of the telescope. If the angle subtended by two distinct objects is less than this angle, the images of the objects cannot be resolved by the telescope. The smaller the value of this angle, higher will be there power of the telescope. Thus, the reciprocal of the limit of resolution gives the resolving power of the telescope.

If  $\lambda$  is the wavelength of light,  $D$  the diameter of the telescope objective, and  $\theta$  the angle subtended by the point object at the objective, the limit of resolution of the telescope is given by (Rayleigh's criterion)

$$\theta = \frac{1.22\lambda}{D}$$

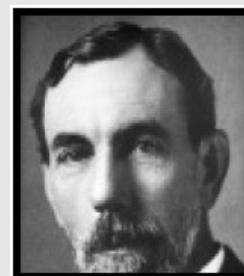
Hence, the resolving power of the telescope.

$$(R.P)_T = \frac{1}{\theta} \frac{D}{1.22\lambda} \quad (15.15)$$

From Eqn. (15.15) it is clear that to get a high resolving power, a telescope with large aperture objective or light of lower wavelength has to be used.

### Lord Rayleigh (1842 – 1919)

Born to the second Baron Rayleigh of Terling place, Witham, in the country of Essex, England, John strutt had very poor health in his childhood. Due to this he had disrupted schooling. But he had the good luck of having Edward Rath and Stokes as his teachers. As a result, he passed his tripos examination in 1865 as senior Wrangler and become the first recipient of Smiths prize.



In addition to the discovery of Argon, for which he was awarded Nobel prize (1904), Rayleigh did extensive work in the fields of hydrodynamics, thermodynamics, optics and mathematics. His travelling wave theory, which suggested that elastic waves can be guided by a surface, paved way for researches in seismology and electronic signal processing. During the later years of his life, he also showed interest in psychiatry research. Lunar feature crater Rayleigh and planetary feature crater Rayleigh on Mars are a tribute to his contributions.

#### Example 15.3

A telescope of aperture 3 cm is focussed on a window at 80 metre distance fitted with a wire mesh of spacing 2 mm. Will the telescope be able to observe the wire mesh? Mean wavelength of light  $\lambda = 5.5 \times 10^{-7}$  m.

**Solution :**

Given  $\lambda = 5.5 \times 10^{-7} \text{ m}$  and  $D = 3 \text{ cm} = 3 \times 10^{-2} \text{ m}$

Therefore, the limit of resolution

$$\theta = \frac{1.22\lambda}{D} = \frac{1.22 \times 5.5 \times 10^{-7} \text{ m}}{3 \times 10^{-2} \text{ m}} = 2.236 \times 10^{-5} \text{ rad}$$

The telescope will be able to resolve the wire mesh, if the angle subtended by it on the objective is equal to or greater than  $\theta$ , the limit of resolution. The angle subtended by the wire mesh on the objective

$$\alpha = \frac{\text{spacing of wire mesh}}{\text{distance of the objective from the wire mesh}}$$

$$= \frac{2 \text{ mm}}{80 \text{ m}} = \frac{2 \times 10^{-3}}{80 \text{ m}} = 2.5 \times 10^{-5} \text{ rad.}$$

As the angle  $2.5 \times 10^{-5}$  radian exceeds the limit of a resolution ( $= 2.236 \times 10^{-5}$  radian), the telescope will be able to observe the wire mesh.

**15.3.2 Resolving Power of a Microscope**

The resolving power of a microscope represents its ability to form separate images of two objects situated very close to each other. The resolving power of a microscope is measured in terms of the smallest linear separation between the two objects which can just be seen through the microscope as separate. This smallest linear separation between two objects is called **the limit of resolution of the microscope**.

The smaller the value of linear separation, the higher will be the resolving power of the microscope. Thus, **the reciprocal of the limit of resolution gives the resolving power of the microscope**.

If  $\lambda$  is the wavelength of light used to illuminate the object,  $\theta$  is the half angle of the cone of light (it is the extent to which an incident beam is converging or diverging) from the point object at the eye and  $n$  is the refractive index of the medium between the object and the objective, the limit of resolution of the microscope is given by

$$d = \frac{\lambda}{2n \sin \theta} \quad (15.16)$$

Thus the resolving power of microscope will be

$$(\text{R.P})_m = \frac{2n \sin \theta}{\lambda} \quad (15.17)$$

The expression  $2n \sin \theta$  is called Numerical Aperture (N.A) The highest value of N.A of the objective obtainable in practice is 1.6, and for the eye, N.A is 0.004.

It is clear from Eqn.(15.17) that the resolving power of a microscope can be increased by increasing the numerical aperture and decreasing the wavelength of the light used to illuminate the object. That is why ultraviolet microscopes and an electron microscope have a very high resolving power.

### Intext Questions 15.3

1. How can the resolving power of a telescope be improved?
2. What is the relationship between the limit of resolution and the resolving power of the eye?
3. If the wavelength of the light used to illuminate the object is increased, what will be the effect on the limit of resolution of the microscope?
4. If a telescope objective is made of larger diameter and light of shorter wavelength is used, how would the resolving power change?

### WHAT YOU HAVE LEARNT

- The angle subtended by an object at the human eye is called as the visual angle.
- The angular magnification or magnifying power of a microscope is the ratio of the angle subtended by the image at the eye to the angle subtended by the object when both are placed at the near point.
- Linear magnification is defined as the ratio of the size of the image to the size of the object.
- The magnifying power of a simple microscope is  $M = 1 + \frac{D}{f}$ , where  $D$  is least distance of distinct vision and  $f$  is focal length of the lens.
- In a compound microscope, unlike the simple microscope, magnification takes place at two stages. There is an eye piece and an objective both having short focal lengths. But the focal length of the objective is comparatively shorter than that of the eye piece.
- The magnifying power of a compound microscope is given as  $M = m_o \times m_e$

But  $m_e = 1 + \frac{D}{f}$ . Therefore

$$M = \frac{v_o}{u_o} \left( 1 + \frac{D}{f} \right)$$

where  $v_o$  is distance between the image and the objective,  $u_o$  is object distance from the objective,  $D$  is the least distance of distinct vision ( $= 25\text{cm}$ ) and  $f_e$  is focal length of the eye-piece.

- Telescope is used to see the distant objects which subtend very small visual angle at the eye. The use of a telescope increases the visual angle at the eye. The far-off object appears to be closer to the eye and can be seen easily.
- Two types of telescopes are used (i) Refracting and (ii) Reflecting.
- The objective of the refracting telescope is a converging lens. But the objective in a reflecting telescope is a spherical mirror of large focal length. There are several advantages of reflecting telescope over a refracting telescope.

The magnifying power of a telescope is  $M = f_o/f_e$

where  $f_o$  is focal length of the objective and  $f_e$  is focal length of the eyepiece.

### TERMINAL EXERCISE

1. What is the difference between simple and compound microscopes? Derive an expression for the magnification of a compound microscope.
2. Distinguish between the refracting telescope and reflecting telescope. Draw a ray diagram for the Newton's telescope.
3. Derive an equation for the magnifying power of a refracting telescope.
4. What do you mean by the least distance of distinct vision? What is its value for a normal eye?
5. Can we photograph the image formed by a compound microscope? Explain your answer.
6. Define the resolving power of an optical instrument. What is the value of limit of resolution for a normal eye?
7. What are the main differences in the design of a compound microscope and a terrestrial telescope?
8. The eyepiece of a telescope has a focal length of 10 cm. The distance between the objective and eye piece is 2.1 m. What is the angular magnification of the telescope?
9. The image formed by a microscope objective of focal length 4 mm is 18 cm from its second focal point. The eyepiece has a focal length of 3.125 cm. What is the magnification of the microscope?
10. The objective of a telescope has a diameter three times that of a second telescope. How much more amount of light is gathered by the first telescope as compared to the second?

### ANSWERS TO INTEXT QUESTIONS

#### 15.1

1. Image formed by a simple microscope is virtual erect and magnified. whereas the image formed by a compound microscope is real, inverted and magnified.

2. Magnifying power is the ratio of the angle subtended by the image at eye piece to the angle subtended by the object placed at the near point. Magnification is the ratio of the size of image to the size of object.
3.  $M = 11, M_e = 1 + \frac{D}{f}$ . Putting  $D = 25$  cm, we get  $f = 2.5$  cm.
4. If you choose the lens with 4 cm focal length, the magnifying power will be high because  $m = \frac{f_o}{f_e}$ .
5. The magnifying power of a compound microscope is given by  $M = \frac{L}{f_o} \left( 1 + \frac{D}{f_e} \right)$   
Obviously,  $M$  will have a large value, if both  $f_o$  and  $f_e$  are small.

### 15.2

1. (a) Objective of large focal length increases the magnifying power of the telescope.  
(b) Magnification is reduced by increasing the focal length of eyepiece.
2. Magnification  $m = \frac{f_o}{f_e} = \frac{50\text{ cm}}{2\text{ cm}} = 25$
3. The objectives of a reflecting telescope is a spherical mirror of large focal length instead of converging lens as in a refracting telescope.
4. A telescope is said to be in normal adjustment, if the final image is formed at infinity.
5. No

### 15.3

1. By taking a large aperture or by using a light of lower wavelength.
2. The limit of resolution of an eye is inversely proportional to its resolving power. Limit of resolution will also be increased.
3. Since resolving power of telescope is given by  $RP = \frac{D}{1.22\lambda}$ , it would increase.

## ANSWERS TO TERMINAL EXERCISE

8. 21                      9. 400                      10. 9 times.



## INTRODUCTION

Reflection, refraction, dispersion and scattering of light can be explained by assuming light travels in a straight line. However, this concept failed to explain bending of light around corners and superposition. Christian Huygens, postulated that light is a wave and through experimental observations explained interference, diffraction and polarization.

## OBJECTIVES

After studying this lesson, you should be able to

- state Huygens' principle and apply it to explain wave propagation;
- explain the phenomena of interference and diffraction of light;
- explain diffraction of light by a single-slit;
- show that polarisation of light established its wave nature ;
- derive Brewster's law;
- explain dispersion of light;
- derive relation between the angle of deviation ( $\delta$ ), angle of prism ( $A$ ) and refractive Index of the material of the prism ( $\mu$ );
- relate the refractive index with wavelength and explain dispersion through a prism;
- explain formation of primary and secondary rainbows;
- explain scattering of light and list its applications;
- explain Raman effect.

## 16.1 HUYGENS' PRINCIPLE

Huygens' postulated that light is a wave, which travels through a hypothetical medium called ether. The vibrations from the source of light propagate in the form of waves and the energy carried by them is distributed equally in all directions. This can be correlated to water waves Fig. 16.1.

The Huygens' principle states that

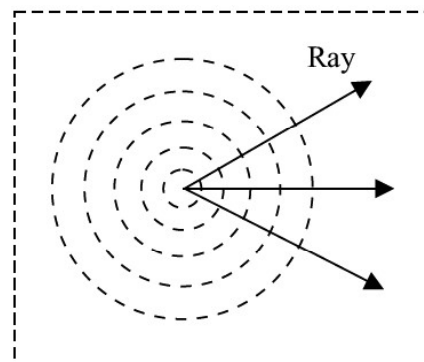


Fig. 16.1 : Circular Wavefronts on the surface of water

- Each point on a wavefront becomes a source of secondary disturbance which spreads out in the medium.
- The position of wavefront at any later instant may be obtained by drawing a forward common envelop to all these secondary wavelets at that instant.
- In an isotropic medium, the energy carried by waves is transmitted equally in all directions.
- If the initial shape, position, the direction of motion and the speed of the wavefront is known, its position at a later instant can be ascertained by geometrical construction. Note that the wavefront does not travel in the backward direction.

To visualise Huygens' construction, you may imagine a point source at the centre of a hollow sphere. The outer surface of this sphere acts as a primary wavefront. If this sphere is enclosed by another hollow sphere of larger radius, the outer surface of the second hollow sphere will act as a secondary wavefront. If the second sphere is further enclosed by another sphere of still bigger radius, the surface of the outermost (third) sphere becomes secondary wavefront and the middle (second) sphere acts as the primary wavefront. In two dimensions, the primary and secondary wavefronts appear as concentric circles.

### 16.1.1 Propagation of Waves

Now let us use Huygens' principle to describe the propagation of light waves in the form of propagation of wavefronts. Fig. 16.2 shows the shape and location of a plane wavefront  $AB$  at the time  $t = 0$ . You should note that the line  $AB$  lies in a plane perpendicular to the plane of the paper. Dots represented by  $a, b, c, d$  on the wavefront  $AB$  are the sources of secondary wavelets. All these sources emit secondary wavelets at the same time and they all travel with the same speed along the direction of motion of the wavefront  $AB$ . In Fig. 16.2, the circular arcs represent the wavelets emitted from  $a, b, c, d$  taking each point as center. *These wavelets have been obtained by drawing arcs of radius,  $r = vt$ , where  $v$  is the velocity of the wavefront and  $t$  is the time at which we wish to obtain the wavefront. The tangent,  $CD$ , to all these wavelets represents the new wavefront at time  $t = T$ .*

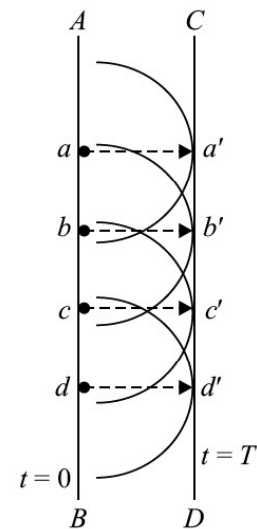


Fig. 16.2 : Construction of a plane wavefront

Let us take another example of Huygens' construction for an expanding circular wavefront. Refer to Fig. 16.3, which indicates a circular wavefront, centred at  $O$ , at time  $t = 0$ . Position  $A, B, C \dots$  represent point sources on this wavefront. Now to draw the wavefront at a later time  $t = T$ , what would you do? You should draw arcs from the points  $A, B, C \dots$ , of radius equal to the speed of the expanding wavefront multiplied by  $T$ .  $A', B', C' \dots$  are arcs indicated in Fig.16.3. These arcs will represent secondary wavelets. The tangents drawn to these arcs will determine the shape and location of the expanding circular wavefront at time  $T$ .

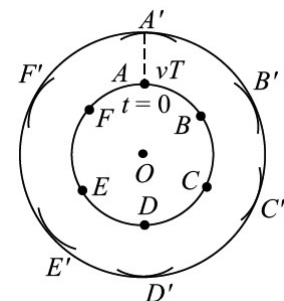


Fig. 16.3 : Construction of circular wavefront using Huygens' Principle



### Intext Questions 16.1

1. What is the relative orientation of a wavefront and the direction of propagation of the wave?
2. A source of secondary disturbance is emitting wavelets at an instant  $t = 0$  s. Calculate the ratio of the radii of wavelets at  $t = 3$  s and  $t = 6$  s.

## 16.2 INTERFERENCE OF LIGHT

*Interference of light refers to redistribution of energy due to superposition of light waves from two coherent sources.* The phenomenon of interference of light was first observed experimentally by Thomas Young in 1802 in his famous double slit experiment. The basic theoretical principle involved in the phenomenon of interference as well as diffraction of light is the *superposition principle*.

### Thomas Young (1773-1829)

Born on 16 June, 1773, Thomas Young will always be known for his study on the human ear, the human eye, how it focuses and on astigmatism. His research on colour blindness led him to the three component theory of colour vision. Working on human ears and eyes, he dedicated much time to the speed of sound and light. He knew that if two sound waves of equal intensity reached the ear  $180^\circ$  out of phase, they cancelled out each other's effect and no sound was heard. It occurred to him that a similar interference effect should be observed with two light beams, if light consisted of waves. This led Young to devise an experiment, now commonly referred to as the Young's double-slit experiment.



In his later years, Young devoted most of his time deciphering the Egyptian hieroglyphics found on the Rosetta stone discovered in the Nile Delta in 1799.

#### 16.2.1 Young's Double Slit Experiment

Young's experimental set up is shown schematically in Fig. 16.4. In his experiment, sunlight was allowed to pass through a pin hole  $S_0$  and then, at some distance away, through two pin holes  $S_1$  and  $S_2$  equidistant from  $S_0$  and close to each other. According to Huygens' wave theory of light, spherical wavefronts would spread out from the pin hole  $S_0$  which get divided into two wavefronts by  $S_1$  and  $S_2$ . If  $S_0$  is illuminated by a monochromatic source of light, such as sodium, it acts as a coherent source, and in-phase waves of equal amplitude from these source superpose as they

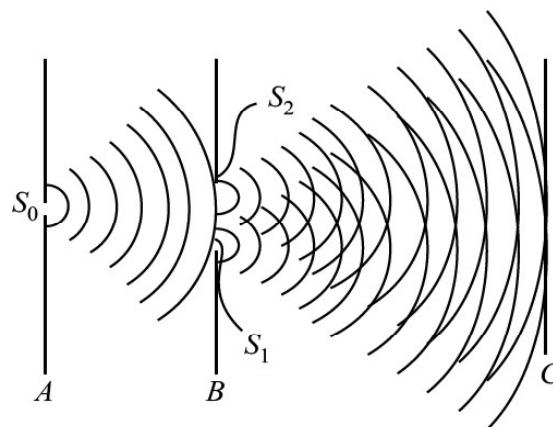


Fig. 16.4 : Schematic arrangement of Young's double-slit experiment

propagate beyond  $S_1S_2$ . As a consequence of superposition (of the two sets of identical waves from  $S_1$  and  $S_2$ ), redistribution of energy takes place and a pattern consisting of alternate bright and dark fringes is produced on the screen such as placed at  $C$ .  $A$  is the position of  $S_0$  and  $B$  is the position of  $S_1$  and  $S_2$  pin holes. Let us now learn the explanation of the observed fringe pattern in the Young's interference experiment.

- (a) **Constructive Interference** : You may recall from the superposition principle that some points on the screen  $C$  will have maximum displacement (or amplitude) because the crests due to one set of waves coincide with the crests due to another set of waves. In other words, at this point, the waves arrive in-phase and hence the total amplitude is much higher than the amplitude of individual waves. The same holds true for the points where the troughs due to one set of waves coincide with the troughs due to another set. Such points will appear bright because the intensity of light wave is proportional to the square of the amplitude. Superposition of waves at these points leads to what is known as **constructive interference**.
- (b) **Destructive Interference** : The points where the crests due to one set of waves coincide with the troughs due to the other set and vice-versa, the total amplitude is zero. It is so because the waves reach these points completely out of phase. Such points appear dark on the screen. These points correspond to **destructive interference**.
- (c) **Intensity of fringes** : To analyse the interference pattern, we calculate the intensity of the bright and dark fringes in the interference pattern for harmonic waves. Refer to Fig. 16.5, which is schematic representation of the geometry of Young's experiment. The phenomenon of interference arises due to superposition of two harmonic waves of same frequency and amplitude but differing in phase. Let the phase difference between these two waves be  $\delta$ . We can write  $y_1$  and  $y_2$ , the displacements at a fixed point  $P$  due to the two waves, as

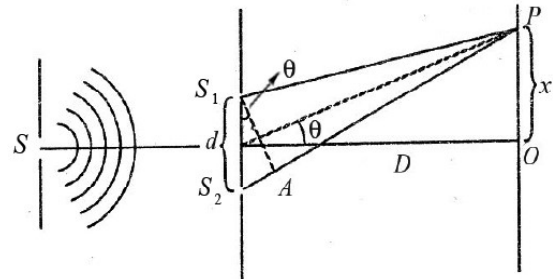


Fig. 16.5 : Geometry of Young's double slit experiment

$$y_1 = a \sin \omega t$$

and 
$$y_2 = a \sin(\omega t + \delta)$$

where  $\delta$  signifies the phase difference between these waves. Note that we have not included the spatial term because we are considering a fixed point in space.

According to the principle of superposition of waves, the resultant displacement is given by

$$\begin{aligned} y &= y_1 + y_2 \\ &= a \sin \omega t + a \sin(\omega t + \delta) \\ &= a[\sin \omega t + a \sin(\omega t + \delta)] \end{aligned}$$

$$\begin{aligned}
 &= 2a \sin\left(\omega t + \frac{\delta}{2}\right) \cos\left(-\frac{\delta}{2}\right) \\
 &= A \sin\left(\omega t + \frac{\delta}{2}\right)
 \end{aligned}$$

where amplitude of the resultant wave is given by

$$A = 2a \cos(\delta/2)$$

The intensity of the resultant wave at point  $P$  can be expressed as

$$\begin{aligned}
 I &\propto A^2 \\
 &\propto 4a^2 \cos^2(\delta/2) \qquad (16.1)
 \end{aligned}$$

To see the dependence of intensity on the phase difference between the two waves, let us consider the following two cases.

**Case 1 :** When the phase difference,  $\delta = 0, 2\pi, 4\pi, \dots, 2n\pi$

$$\begin{aligned}
 I &= 4a^2 \cos^2 0 \\
 &= 4a^2
 \end{aligned}$$

**Case 2 :** When,  $\delta = \pi, 3\pi, 5\pi, \dots, (2n+1)\pi$

$$\begin{aligned}
 I &= 4a^2 \cos^2(\delta/2) \\
 &= 0
 \end{aligned}$$

From these results we can conclude that when phase difference between superposing waves is an integral multiple of  $2\pi$ , the two waves arrive at the screen ‘in-phase’ and the resultant intensity (or the brightness) at those points is more than that due to individual waves (which is equal to  $4a^2$ ). On the other hand, when phase difference between the two superposing waves is an odd multiple of  $\pi$ , the two superposing waves arrive at the screen ‘out of phase’. Such points have zero intensity and appear to be dark on the screen.

**(d) Phase Difference and Path Difference :** It is obvious from the above discussion that to know whether a point on the screen will be bright or dark, we need to know the phase difference between the waves arriving at that point. The phase difference can be expressed in terms of the path difference between the waves during their journey from the sources to a point on the observation screen. You may recall that waves starting from  $S_1$  and  $S_2$  are in phase. Thus, whatever phase difference arises between them at the point  $P$  is because of the different paths travelled by them upto observation point from  $S_1$  and  $S_2$ . From Fig. 16.5, we can write the path difference as

$$\Delta = S_2P - S_1P$$

We know that path difference of one wavelength is equivalent to a phase difference of  $2\pi$ . Thus, the relation between the phase difference  $\delta$  and the path difference  $\Delta$  is

$$\Delta = \left( \frac{\lambda}{2\pi} \right) \delta \quad (16.2)$$

From Eqn. (16.1) we note that bright fringes (corresponding to constructive interference) are observed when the phase difference is  $2n\pi$ . Using this in Eqn. (16.2) we find that the path difference for observing bright fringes is

$$(\Delta)_{\text{bright}} = \left( \frac{\lambda}{2\pi} \right) 2n\pi = n\lambda; \quad n = 0, 1, 2, \dots \quad (16.3)$$

Similarly, for dark fringes, we get

$$\begin{aligned} (\Delta)_{\text{dark}} &= (\lambda/2\pi)(2n+1)\pi \\ &= (2n+1) \frac{\pi}{2}; \quad n = 0, 1, 2, \dots \end{aligned} \quad (16.4)$$

Having obtained expressions for the bright and dark fringes in terms of the path difference and the wavelength of the light used, let us now relate path difference with the geometry of the experiment, i.e., relate  $\Delta$  with the distance  $D$  between the source and the screen, separation between the pin holes ( $d$ ) and the location of the point  $P$  on the screen. From Fig. 16.5 we note that

$$\Delta = S_2P - S_1P = S_2A = d \sin \theta$$

Assuming  $\theta$  to be small, we can write

$$\sin \theta \approx \tan \theta \approx \theta$$

and

$$\sin \theta = x / D$$

Therefore, the expression for path difference can be rewritten as

$$\Delta = d \sin \theta = x \frac{d}{D} \quad (16.5)$$

On substituting Eqn. (16.5) in Eqns. (16.2) and (16.3), we get

$$\frac{d}{D}(x_n)_{\text{bright}} = n\lambda$$

or

$$(x_n)_{\text{bright}} = \frac{n\lambda D}{d}; \quad n = 1, 2, \dots \quad (16.6)$$

and 
$$\frac{d}{D}(x_n)_{\text{dark}} = \left(n + \frac{1}{2}\right)\lambda$$

or 
$$(x_n)_{\text{dark}} = \left(n + \frac{1}{2}\right)\frac{\lambda D}{d}; n = 0, 1, 2, \dots \quad (16.7)$$

Eqns. (16.6) and (16.7) specify the positions of the bright and dark fringes on the screen.

- (e) **Fringe width :** You may now ask: How wide is a bright or a dark fringe? To answer this question, we first determine the location of two consecutive bright (or dark) fringes. Let us first do it for bright fringes. For third and second bright fringes, from Eqn. (16.6), we can write

$$(x_3)_{\text{bright}} = 3\frac{\lambda D}{d}$$

and 
$$(x_2)_{\text{bright}} = 2\frac{\lambda D}{d}$$

Therefore, fringe width,  $\beta$  is given by

$$\beta = (x_3)_{\text{bright}} - (x_2)_{\text{bright}} = \frac{\lambda D}{d} \quad (16.8)$$

You should convince yourself that the fringe width of an interference pattern remains the same for any two consecutive value of  $n$ . Note that fringe width is directly proportional to linear power of wavelength and distance between the source plane and screen and inversely proportional to the distance between the slits. In actual practice, fringes are so fine that we use a magnifying glass to see them.

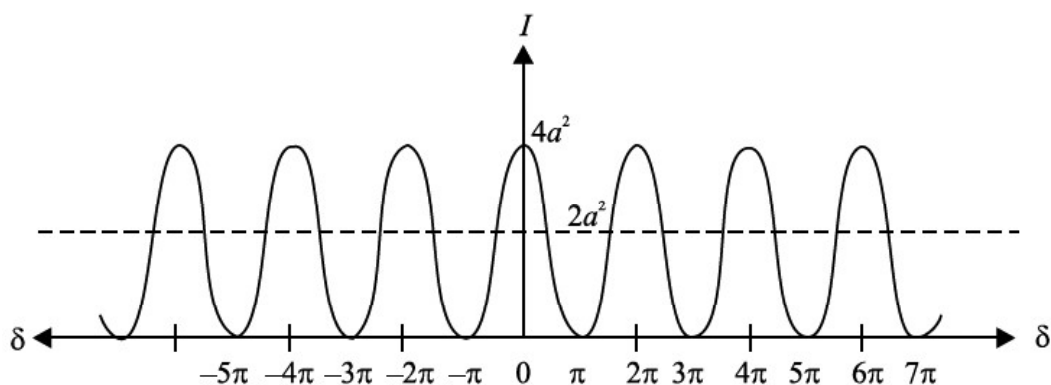


Fig. 16.6 : Intensity distribution in an interference pattern

Next let us learn about the intensity of bright and dark fringes in the interference pattern. We know that when two light waves arrive at a point on the screen out of phase, we get dark fringes. You may ask : Does this phenomenon not violate the law of conservation of

energy because energy carried by two light waves seem to be destroyed? It is not so; the energy conservation principle is not violated in the interference pattern. Actually, the energy which disappears at the dark fringes reappears at the bright fringes. You may note from Eqn. (16.1) that the intensity of the bright fringes is four times the intensity due to an individual wave. Therefore, in an interference fringe pattern, shown in Fig. 16.6, **the energy is redistributed and it varies between  $4a^2$  and zero**. Each beam, acting independently, will contribute  $a^2$  and hence, in the absence of interference, the screen will be uniformly illuminated with intensity  $2a^2$  due to the light coming from two identical sources. This is the average intensity shown by the broken line in Fig. 16.6.

### Intext Questions 16.2

1. On what factors does the resultant displacement at any point in the region of superposition of two waves depend?
2. In Young's experiment, how is the constructive interference produced on the screen?
3. If we replace the pinholes  $S_1$  and  $S_2$  by two incandescent light bulbs, can we still observe the bright and dark fringes on the screen?
4. What are coherent sources? Can our eyes also act as coherent sources?

## 16.3 DIFFRACTION OF LIGHT

The most obvious manifestation of the rectilinear propagation of light is in the formation of shadow. But, if you study formation of shadows carefully, you will find that, as such, these are not sharp at the edges. For example, the law of rectilinear propagation is violated when the light passes through a very narrow aperture or falls on an obstacle of very small dimensions. At the edges of the aperture or the obstacle, light bends into the shadow region and does not propagate along a straight line. ***This bending of light around the edges of an obstacle is known as diffraction.***

Before discussing the phenomenon of diffraction of light in detail, you may like to observe diffraction of light yourself. Here is a simple situation. Look at the street light at night and almost close your eyes. What do you see? The light will appear to streak out from the lamp/tube. This happens due to the diffraction (bending) of light round the corners of your eyelids.

Another way to observe diffraction is to use a handkerchief. Hold it close to your eyes and look at the Sun or a lamp. You will observe circular fringes, which form due to diffraction of light by small apertures formed by crissed-crossed threads.

In the above situations, the dimensions of the diffracting obstacle/aperture are very small. To observe diffraction, either of the following conditions must be satisfied:

- a) *The size of the obstacle or the aperture should be of the order of the wavelength of the incident wave.*
- b) *The separation between the obstacle or aperture and the screen should be considerably larger (a few thousand times) than the size of the obstacle or aperture.*

On the basis of the above observations, it is easy to understand why we normally do not observe diffraction of light and why light appears to travel in a straight line. You know that the wavelength of light is of the order of  $10^{-6}$  m. Therefore, *to observe diffraction of light, we need to have obstacles or aperture having dimensions of this order!*

### 16.3.1 Diffraction at a Single Slit

Let us see how diffraction pattern appears for a simple opening like a single slit. Refer to Fig. 16.7. It shows the experimental arrangement for producing diffraction pattern.  $S$  is a monochromatic source of light. It is placed on the focal plane of a converging lens so that a plane wavefront is incident on a narrow slit. Another converging lens focusses light from different portions of the slit on the observation screen.

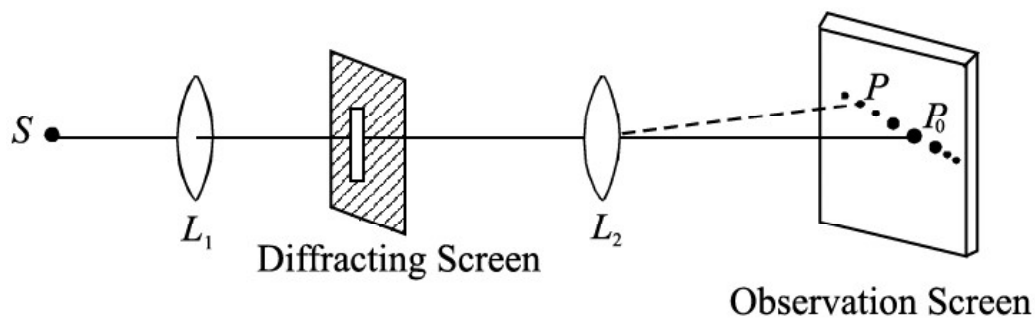


Fig. 16.7 : Schematic representation of single slit diffraction

The salient features of the actual diffraction pattern produced by a single vertical slit from a point source as shown in Fig. 16.8 are :

- A horizontal streak of light along a line normal to the length of the slit.
- The horizontal pattern is a series of bright spots.
- The spot at the centre is the brightest. On either side of this spot, we observe a few more symmetrically situated bright spots of diminishing intensity. The central spot is called *principal maxima* and other spots are called *secondary maxima*.
- The width of the central spot is twice the width of other spots.

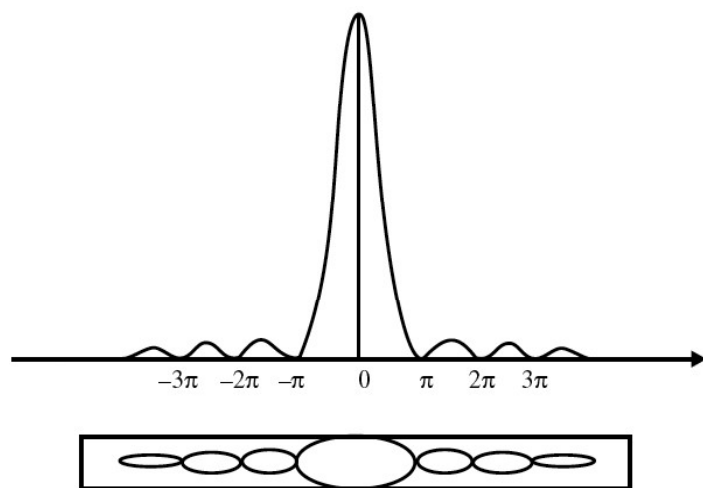


Fig. 16.8 : Observed diffraction pattern single of slit

To understand the theoretical basis of these results, we note that according to Huygens' wave theory, plane wavefronts are incident on the barrier containing the slit. As these wavefronts fall on the barrier, only that part of the wavefront passes through the slit which is incident

on it. This part of the wavefront continues to propagate to the right of the barrier. However, the shape of the wavefront does not remain plane beyond the slit.

Refer to Fig. 16.9 which shows that each point of the aperture such as  $QPR \dots Q'$  form a series of coherent sources of secondary wavelets. In the central part of the wavefront to the right of the barrier, the wavelet emitted from the point  $P$ , say, spreads because of the presence of wavelets on its both sides emitted from the points such as  $Q$  and  $R$ . Since the shape of the wavefront is determined by the tangent to these wavelets, the central part of the wavefront remains plane as it propagates. But for the wavelets emitted from points  $Q$  and  $Q'$  near the edges of the slit, there are no wavelets beyond the edges with which these may superpose. Since the superposition helps to maintain the shape of the wavefront as plane, the absence of such superposing wavelets for the wavelets emitted from the points near the edges allows them to deviate from their plain shape. In other words, the wavelets at the edges tend to spread out. As a result, the plane wavefront incident on a thin aperture of finite size, after passing through it does not remain plane.

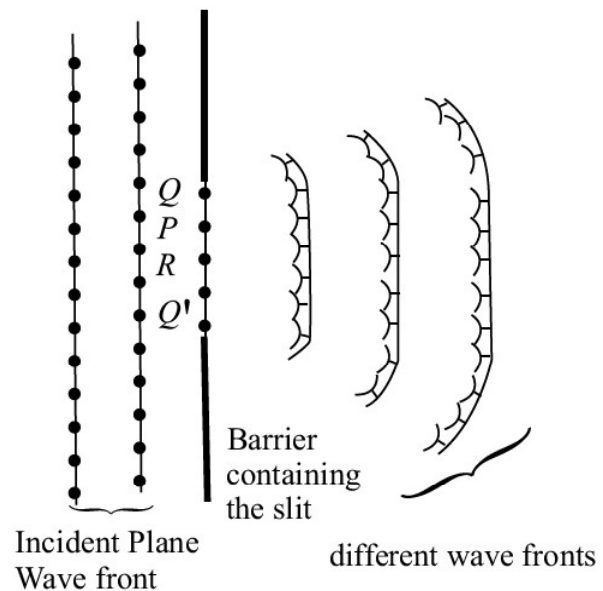


Fig. 16.9 : Huygen's construction for diffraction of light from a narrow slit

To understand the intensity distribution of the single-slit diffraction pattern, we determine the nature of the superposition of waves reaching the screen. In order to apply Huygens' principle, let us divide the width ' $a$ ' of the slit into, say, 100 equal parts. Each of these can be considered as a source of secondary wavelets. The wavelets emanating from these points spread out into the region to the right of the slit. Since the plane wavefront is incident on the slit, initially all points on it are in phase. Therefore, the wavelets emitted by these points are all in phase at the time of leaving the slit. Now let us consider the effect of the superposition of these wavelets at point  $O$  on the screen. The symmetry of the Fig. 16.10 suggests that the wavelets emitted from source of 1 and 100 will reach  $O$  in phase. It is so because both the wavelets travel equal path length. When they started their journey from the respective points on the

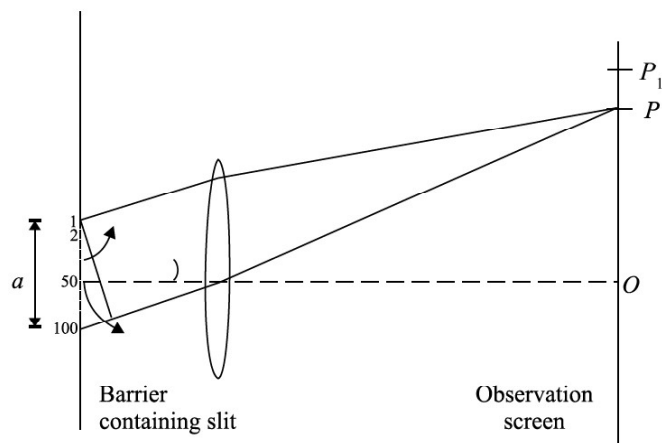


Fig. 16.10 : Schematic representation of single slit diffraction



slit, they were in phase. Hence they arrive at  $O$  in phase and superpose in such a manner as to give resultant amplitude much more than that due to the individual wavelets from the source 1 and 100. Similarly, for each wavelet from source 2 to 50, we have a corresponding wavelet from the source 99 to 51 which will produce constructive interference causing enhancement in intensity at the center  $O$ . Thus the point  $O$  will appear bright on the screen.

Now let us consider an off-axis point  $P$  on the observation screen. Suppose that point  $P$  is such that the path difference between the extreme points i.e. sources 1 and 100 is equal to  $\lambda$ . Thus the path difference between the wavelets from source 1 and 51 will be nearly equal to  $(\lambda/2)$ .

You may recall from the interference of light that the waves coming from the sources 1 and 51 will arrive at  $P$  out of phase and give rise to destructive interference. Similarly, wavelets from the sources 2 and 52 and all such pair of wavelets will give rise to destructive interference at the point  $P$ . Therefore, we will have minimum intensity at point  $P$ . Similarly, we will get minimum intensity for other points for which the path difference between the source edges is equal to  $2\lambda$ . We can imagine that the slit is divided into four equal points and we can, by similar pairing of 1 and 26, 2 and 27, ... show that first and second quarters have a path difference of  $\lambda/2$  and cancel each other. Third and fourth quarters cancel each other by the same argument so that the resultant intensity will be minimum, and so on. *We can therefore conclude that when the path difference between the extreme waves diffracted by the extreme points in a particular direction is an integral multiple of  $\lambda$ , the resultant diffracted intensity in that direction will be zero.*

Let us now find intensity at a point  $P'$  which lies between the points  $P$  and  $P_1$  (Fig.16.10) and the *path difference between waves diffracted from extreme points is  $3\lambda/2$* . We divide the wavefront at the slit into 3 equal parts. In such a situation, secondary wavelets from the corresponding sources of two parts will have a path difference of  $\lambda/2$  when they reach the point  $P$  and cancel each other. However, wavelets from the third portion of the wavefront will all contribute constructively (presuming that practically the path difference for wavelets from this part is zero) and produce brightness at  $P'$ . Since only one third of the wavefront contributes towards the intensity at  $P'$  as compared to  $O$ , where the whole wavefront contributes, the intensity at  $P'$  is considerably less than that of the intensity at  $O$ . The point  $P'$  and all other similar points constitute secondary maxima.

However, you must note here that this is only a qualitative and simplified explanation of the diffraction at a single slit. You will study more rigorous analysis of this phenomena when you pursue higher studies in Physics.

### Intext Questions 16.3

1. Does the phenomenon of diffraction show that the light does not travel along a straight line path?
2. Distinguish between interference and diffraction of light.
3. Why are the intensity of the principal maximum and the secondary maxima of a single slit diffraction not the same?

## 16.4 POLARISATION OF LIGHT

In the previous two sections of this lesson, you learnt about the phenomena of interference and diffraction of light. While discussing these phenomena, we did not bother to know the nature of light waves; whether these were longitudinal or transverse. However, polarisation of light conclusively established that light is a transverse wave.

To understand the phenomenon of polarisation, you can perform a simple activity.

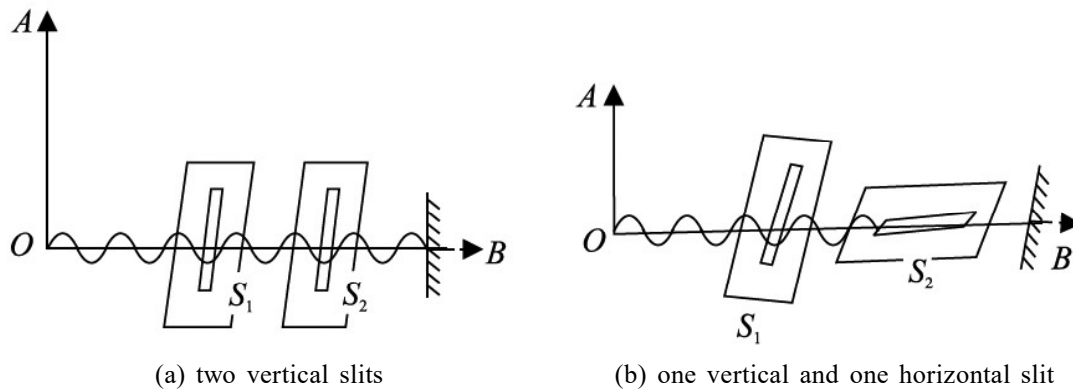


Fig. 16.11 : Transverse wave on a rope passing through

Take two card boards having narrow vertical slits  $S_1$  and  $S_2$  and hold them parallel to each other. Pass a length of a string through the two slits, fix its one end and hold the other in your hand. Now move your hand up and down and sideways to generate waves in all directions. You will see that the waves passing through the vertical slit  $S_1$  will also pass through  $S_2$ , as shown in Fig. 16.11 (a). Repeat the experiment by making the slit  $S_2$  horizontal. You will see no waves beyond  $S_2$ . It means that waves passing through  $S_1$  cannot pass through the horizontal slit  $S_2$ . This is because the vibrations in the wave are in a plane at right angles to the slits  $S_2$ , as shown in Fig. 16.11 (b).

This activity can be repeated for light by placing a source of light at  $O$  and replacing the slits by two polaroids. You will see light in case (a) only. This shows that light has vibrations confined to a plane. It is said to be linearly polarised or plane polarised after passing through the first polaroid Fig. 16.12.

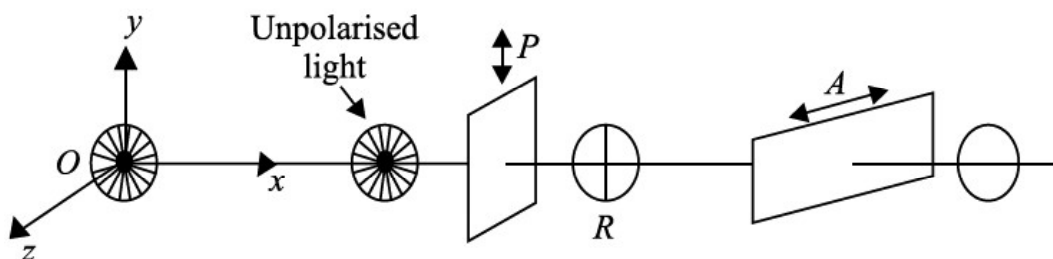


Fig. 16.12 : Schematics of the apparatus for observing polarisation of light

When an unpolarised light falls on glass, water or any other transparent material, the reflected light is, in general, partially plane polarised. Fig. 16.13 shows unpolarised light  $AO$  incident on a glass plate. The reflected light is shown by  $OR$  and the transmitted wave

by  $OT$ . When the light is incident at polarising angle, the polarisation is complete. At this angle, the reflected and transmitted rays are at right angles to each other.

The polarising angle depends on the refractive index of the material of glass plate on which the (unpolarised) beam of light is incident. The relation between  $r$  and  $i_p$  is obtained by using Snell's law (refer Fig. 16.13):

$$\mu = \frac{\sin i_p}{\sin r} = \frac{\sin i_p}{\sin (90 - i_p)}$$

$$= \frac{\sin i_p}{\cos i_p} = \tan i_p$$

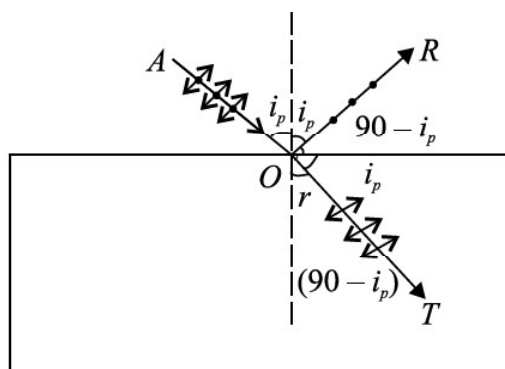


Fig. 16.13 : Polarisation of reflected and refracted light

This is known as **Brewster's law**. It implies that polarising angle  $i_p$  depends on the refractive index of the material. For air water interface,  $i_p = 53^\circ$ . It means that when the sun is  $37^\circ$  above the horizontal, the light reflected from a calm pond or lake will be completely linearly polarised. Brewster's law has many applications in daily life. Glare caused by the light reflected from a smooth surface can be reduced by using polarising materials called polaroids, which are made from tiny crystals of quinine iodosulphate; all lined up in the same direction in a sheet of nitro cellulose. Such crystals (called dichoric) transmit light in one specific plane and absorb those in a perpendicular plan. Thus, polaroid coatings on sunglasses reduce glare by absorbing a component of the polarized light. Polaroid discs are used in photography as 'filters' in front of camera lens and facilitate details which would otherwise be hidden by glare. Polarimeters are used in sugar industry for quality control.

### Intext Questions 16.4

1. Polarisation of light is the surest evidence that light is a transverse wave. Justify.
2. Is it correct to say that the direction of motion of a wave may not lie in the plane of polarisation?
3. Suppose a beam of unpolarised light is incident on a set of two *polaroids*. If you want to block light completely with the help of these polaroids, what should be the angle between the transmission axes of these polaroids?
4. Do sound waves in air exhibit polarization?

## 16.5 DISPERSION OF LIGHT

In the previous lesson you have learnt about reflection, refraction and total internal reflection of light. You have also learnt about image formation by mirrors and lenses and their uses in daily life. When a narrow beam of ordinary light is refracted by a prism, we see colour bands. This phenomenon has to be other than reflection or refraction. *The splitting of white light into its constituent colours or wavelengths by a medium is called dispersion.* In this lesson, you will study about this phenomenon. A beautiful manifestation of this

phenomenon in nature is in the form of rainbow. You will also learn in this lesson about the phenomenon of scattering of light, which gives sky its blue colour and the sun red colour at sunrise and sunset.

Natural phenomena like rings around planets (halos) and formation of rainbow etc. cannot be explained by the rectilinear propagation of light. To understand such events, light is considered as having wave nature. As you know, light waves are transverse electromagnetic waves which propagate with speed  $3 \times 10^8 \text{ ms}^{-1}$  in vacuum. Of the wide range of electromagnetic spectrum, the visible light forms only a small part. Sunlight consists of seven different wavelengths corresponding to seven colours. Thus, colours may be identified with their wavelengths. You have already learnt that the speed and wavelength of waves change when they travel from one medium to another. The speed of light waves and their corresponding wavelengths also change with the change in the medium. The speed of a wave having a certain wavelength becomes less than its speed in free space when it enters an optically denser medium.

The refractive index  $\mu$  has been defined as the ratio of the speed of light in vacuum to the speed of light in the medium. It means that the refractive index of a given medium will be different for waves having wavelengths  $3.8 \times 10^{-7} \text{ m}$  and  $5.8 \times 10^{-7} \text{ m}$  because these waves travel with different speeds in the same medium. This **variation of the refractive index of a material with wavelength is known as dispersion**. This phenomenon is different from refraction. In free space and even in air, the speeds of all waves of the visible light are the same. So, they are not separated. (Such a medium is called a non dispersive medium.) But in an optically denser medium, the component wavelengths (colours) travel with different speeds and therefore get separated. Such a medium is called dispersive medium. Does this suggest that light will exhibit dispersion whenever it passes through an optically denser medium. Let us learn about it now.

### 16.5.1 Dispersion through a Prism

The separation of colours by a medium is not a sufficient condition to observe dispersion of light. These colours must be widely separated and should not mix up again after emerging from the dispersing medium. A glass slab Fig. 16.14 is not suitable for observing dispersion as the rays of the emergent beam are very close and parallel to the incident beam

Newton used a prism to demonstrate dispersion of light. Refer to Fig. 16.15. White light from a slit falls on the face  $AB$  of the prism and light emerging from face  $AC$  is seen to split into different colours. Coloured patches can be seen on a screen. The face  $AC$  increases the separation between the rays refracted at the face  $AB$ . The incident white light  $PQ$  thus splits up into its component seven colours : Violet, indigo, blue, green, yellow, orange and red (VIBGYOR). The wavelengths travelling with different speeds are refracted through different angles and are thus separated. *This splitting of white light into component colours is known as dispersion.  $MR$  and  $MV$  correspond to the red and violet light respectively. These colours on the screen produce the spectrum.*

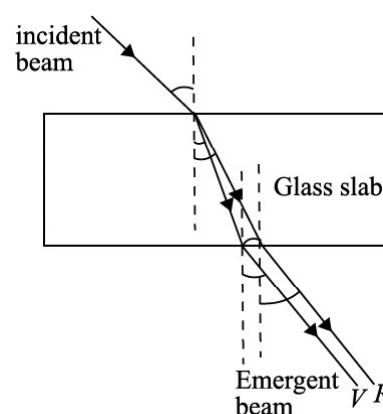


Fig. 16.14 : Passage of light through a glass slab

The bending of the original beam  $PQN$  along  $MR$  and  $MV$  etc. is known as *deviation*. The angle between the emergent ray and the incident ray is known as the **angle of deviation**. Thus  $\delta_v$  and  $\delta_r$  represent the angles of deviation for violet light and red light, respectively.

Read the following example carefully to fix the ideas on variation of the refractive index with the wavelength of light.

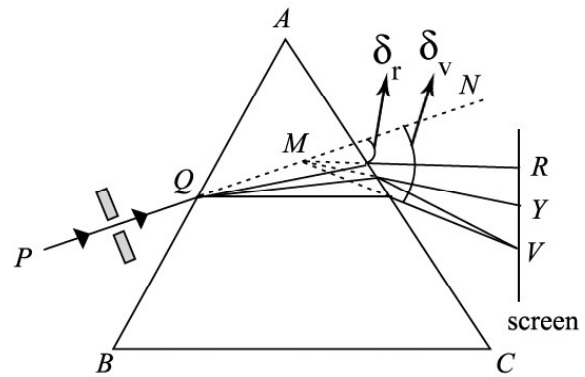


Fig. 16.15 : Dispersion of light by a prism

**Example 16.1**

A beam of light of average wavelength 600 nm, on entering a glass prism, splits into three coloured beams of wavelengths 384 nm, 589 nm and 760 nm respectively. Determine the refractive indices of the material of the prism for these wavelengths.

**Solution :**

The refractive index of the material of the prism is given by

$$\mu = \frac{c}{v}$$

where  $c$  is speed of light in vacuum, and  $v$  is speed of light in the medium (prism).

Since velocity of a wave is product of frequency and wavelength, we can write

$$c = v\lambda_a \quad \text{and} \quad v = v\lambda_m$$

where  $\lambda_a$  and  $\lambda_m$  are the wavelengths in air and medium respectively and  $v$  is the frequency of light waves. Thus

$$\mu = \frac{v\lambda_a}{v\lambda_m} = \frac{\lambda_a}{\lambda_m}$$

For 384 nm wavelength, the refractive index is

$$\mu_1 = \frac{600 \times 10^{-9} \text{ m}}{384 \times 10^{-9} \text{ m}} = 1.56$$

For wave length of 589 nm :

$$\mu_2 = \frac{600 \times 10^{-9} \text{ m}}{58.9 \times 10^{-9} \text{ m}} = 1.02$$

and for 760 nm wavelength :

$$\mu_3 = \frac{600 \times 10^{-9} \text{ m}}{760 \times 10^{-9} \text{ m}} = 0.8$$

We have seen that the refractive index of a material depends on

- the nature of the material, and
- the wavelength of light.

An interesting outcome of the above example is that the variation in wavelength ( $\Delta\lambda = \lambda_2 - \lambda_1$ ) produces variation in the refractive index ( $\Delta\mu = \mu_2 - \mu_1$ ). The ratio  $\frac{\Delta\mu}{\Delta\lambda}$  is known as the spectral *dispersive power of the material of prism*.

### 16.5.2 The Angle of Deviation

We would now establish the relation between the angle of incidence  $i$ , the angle of deviation  $\delta$  and the angle of prism  $A$ . Let us consider that a monochromatic beam of light  $PQ$  is incident on the face  $AB$  of the principal section of the prism  $ABC$  Fig.16.16. On refraction, it goes along  $QR$  inside the prism and emerges along  $RS$  from face  $AC$ . Let  $\angle A = \angle BAC$  be the refracting angle of the prism. We draw normals  $NQ$  and  $MR$  on the faces  $AB$  and  $AC$ , respectively and produce them backward to meet at  $O$ . Then you can easily convince yourself that  $\angle NQP = \angle i$ ,  $\angle MRS = \angle e$ ,  $\angle RQO = \angle r_1$ , and  $\angle QRO = \angle r_2$  are the angle of incidence, the angle of emergence and the angle of refraction at the faces  $AB$  and  $AC$ , respectively. The angle between the emergent ray  $RS$  and the incident ray  $PQ$  at  $D$  is known as the angle of deviation ( $\delta$ ).

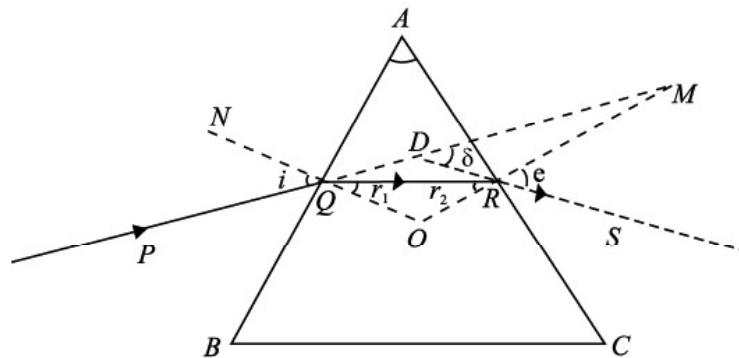


Fig. 16.16 : Refraction through a prism

Since  $\angle MDR = \angle \delta$ , as it is the external angle of the triangle  $QDR$ , we can write

$$\begin{aligned}\angle \delta &= \angle DQR + \angle DRQ \\ &= (\angle i - \angle r_1) + (\angle e - \angle r_2) \\ \text{or} \quad \angle \delta &= (\angle i + \angle e) - (\angle r_1 + \angle r_2) \quad (16.9)\end{aligned}$$

You may recall that the sum of the internal angles of a quadrilateral is equal to  $360^\circ$ . In the quadrilateral  $AQOR$ ,  $\angle AQO = \angle ARO = 90^\circ$ , since  $NQ$  and  $MR$  are normals on faces  $AB$  and  $AC$ , respectively. Therefore

$$\begin{aligned}\angle QAR + \angle QOR &= 180^\circ \\ \text{or} \quad \angle A + \angle QOR &= 180^\circ \quad (16.10)\end{aligned}$$

But in  $\Delta QOR$

$$\begin{aligned}\angle OQR + \angle QRO + \angle QOR &= 180^\circ \\ \text{or} \quad \angle r_1 + \angle r_2 + \angle QOR &= 180^\circ \quad (16.11)\end{aligned}$$

On comparing Eqns. (16.10) and (16.11), we have

$$\angle r_1 + \angle r_2 = \angle A \quad (16.12)$$

Combining this result with Eqn. (21.1), we have

$$\angle \delta = \angle i + \angle e - \angle A$$

or

$$\angle i + \angle e = \angle A + \angle \delta \quad (16.13)$$

### Angle of Minimum Deviation

If we vary the angle of incidence  $i$ , the angle of deviation  $\delta$  also changes; it becomes minimum for a certain value of  $i$  and again starts increasing as  $i$  increases further Fig. 16.17. The minimum value of the angle of deviation is called angle of minimum deviation ( $\delta_m$ ). It depends on the material of the prism and the wavelength of light used. In fact, one angle of deviation may be obtained corresponding to two values of the angles of incidence. Using the principle of reversibility of light, we find that the second value of angle of incidence corresponds to the angle of emergence ( $e$ ). In the minimum deviation position, there is only one value of the angle of incidence. So we have

$$\angle e = \angle i$$

Using this fact in Eqn. (16.13) and replacing  $\delta$  by  $\delta_m$ , we have

$$\angle i = \frac{\angle A + \angle \delta_m}{2} \quad (16.14)$$

Applying the principle of reversibility of light rays and under the condition  $\angle e = \angle i$ , we can write  $\angle r_1 = \angle r_2 = \angle r$ , say

On substituting this result in Eqn. (16.12), we get

$$\angle r = \frac{\angle A}{2} \quad (16.15)$$

The light beam inside the prism, under the condition of minimum deviation, passes symmetrically through the prism and is parallel to its base. The refractive index of the material of the prism is therefore given by

$$\mu = \frac{\sin i}{\sin r} = \frac{\sin\left(\frac{A + \delta_m}{2}\right)}{\sin\frac{A}{2}} \quad (16.16)$$

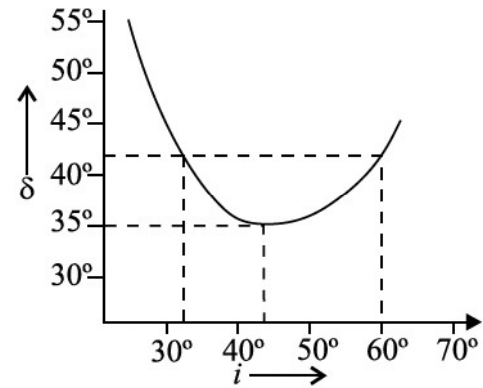


Fig. 16.17 : Plot between angle of incidence  $i$  and angle of deviation  $\delta$

The refractive index  $\mu$  can be calculated using Eqn. (16.16) for a monochromatic or a polychromatic beam of light. The value of  $\delta_m$  is different for different colours. It gives a unique value of the angle of incidence and the emergent beam is brightest for this incidence.

For a prism of small angle  $A$ , keeping  $i$  and  $r$  small, we can write

$$\sin i = i, \sin r = r, \text{ and } \sin e = e$$

Hence

$$\mu = \frac{\sin i}{\sin r_1} = \frac{i}{r_1} \text{ or } i = \mu r_1$$

Also

$$\mu = \frac{\sin e}{\sin r_2} = \frac{e}{r_2} \text{ or } e = \mu r_2$$

Therefore,

$$\angle i + \angle e = \mu (\angle r_1 + \angle r_2)$$

Using this result in Eqns. (16.17) and (16.18), we get

$$\mu \angle A = \angle A + \angle \delta$$

or

$$\angle \delta = (\mu - 1) \angle A \quad (16.17)$$

We know that  $\mu$  depends on the wavelength of light. So deviation will also depend on the wavelength of light. That is why  $\delta_v$  is different from  $\delta_R$ . Since the velocity of the red light is more than that of the violet light in glass, the deviation of the red light would be less as compared to that of the violet light.

$$\delta_v > \delta_R$$

This implies that  $\mu_v > \mu_R$ . This change in the refractive index of the material with the wavelength of light is responsible for dispersion phenomenon.

### 16.5.3 Angular Dispersion and Dispersive Power

The difference between the angles of deviation for any two wavelengths (colours) is known as the angular dispersion for those wavelengths. The angular dispersion between the red and violet wavelengths is  $\delta_v - \delta_R$ . In the visible part of the spectrum, the wavelength of the yellow colour is nearly the average wavelength of the spectrum. The deviation for this colour  $\delta_Y$  may, therefore, be taken as the average of all deviations.

**The ratio of the angular dispersion to the mean deviation is taken as the dispersive power ( $\omega$ ) of the material of the prism :**

$$\omega = \frac{\delta_v - \delta_R}{\delta_Y}$$

We can express this result in terms of the refractive indices using Eqn.(16.17):

$$\omega = \frac{(\mu_v - 1)\angle A - (\mu_R - 1)\angle A}{(\mu_v - 1)\angle A}$$



$$= \frac{\mu_v - \mu_R}{\mu_Y - 1} = \frac{\Delta\mu}{\mu - 1} \quad (16.18)$$

**Example 16.2**

The refracting angle of a prism is  $30'$  and its refractive index is 1.6. Calculate the deviation caused by the prism.

**Solution :**

We know that  $\delta = (\mu - 1) \angle A$

On substituting the given data, we get

$$\delta = (1.6 - 1) \times \frac{1^\circ}{2} = \frac{0.6}{2} = 0.3^\circ = 18'$$

**Example 16.3**

For a prism of angle  $A$ , the angle of minimum deviation is  $A/2$ . Calculate its refractive index, when a monochromatic light is used. Given  $A = 60^\circ$

**Solution :**

The refractive index is given by

$$\mu = \frac{\sin\left(\frac{A + \delta_m}{2}\right)}{\sin(A/2)}$$

Now  $\delta_m = A/2$  so that

$$\mu = \frac{\sin\left(\frac{A + A/2}{2}\right)}{\sin(A/2)} = \frac{\sin\left(\frac{3}{4}A\right)}{\sin\left(\frac{A}{2}\right)} = \frac{\sin\left(\frac{3}{4}A\right)}{\sin\left(\frac{A}{2}\right)} = \sqrt{2} = 1.4$$

**Intext Questions 16.5**

1. Most ordinary gases do not show dispersion with visible light. Why?
2. With your knowledge about the relative values of  $\mu$  for the component colours of white light, state which colour is deviated more from its original direction?
3. Does dispersion depend on the size and angle of the prism?
4. Calculate the refractive index of an equilateral prism if the angle of minimum deviation is equal to the angle of the prism.

## Rainbow formation

Dispersion of sunlight through suspended water drops in air produces a spectacular effect in nature in the form of rainbow on a rainy day. With Sun at our back, we can see a brighter and another fainter rainbow. The brighter one is called the **primary** rainbow and the other one is said to be **secondary rainbow**. Sometimes we see only one **rainbow**. The bows are in the form of coloured arcs whose common centre lies at the line joining the Sun and our eye. Rainbow can also be seen in a fountain of water in the evening or morning when the sun rays are incident on the water drops at a definite angle.

### Primary Rainbow

The primary rainbow is formed by two refractions and a single internal reflection of sunlight in a water drop. (See Fig. 16.18(a)). Descartes explained that rainbow is seen through the rays which have suffered minimum deviation. Parallel rays from the Sun suffering deviation of  $137^{\circ}29'$  or making an angle of  $42^{\circ}31'$  at the eye with the incident ray, after emerging from the water drop, produce bright shining colours in the bow. Dispersion by water causes different colours (red to violet) to make their own arcs which lie within a cone of  $43^{\circ}$  for red and  $41^{\circ}$  for violet rays on the outer and inner sides of the bow (Fig. 16.18 (b)).

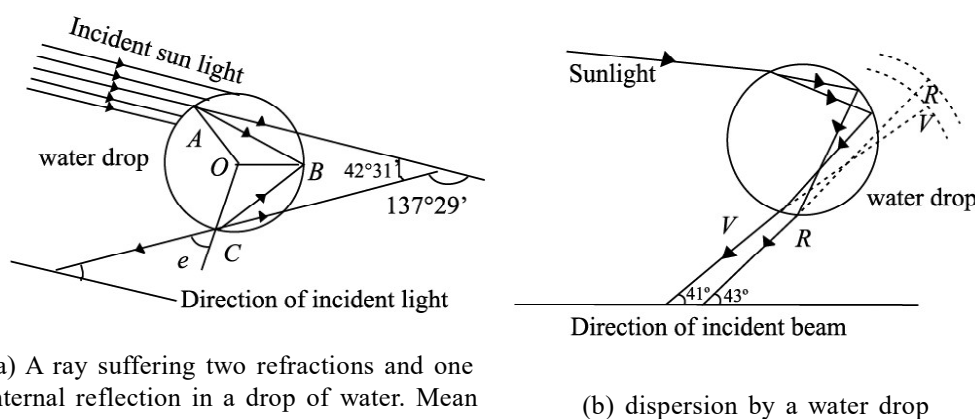


Fig. 16.18

### Secondary Rainbow

The secondary rainbow is formed by two refractions and two internal reflections of light on the water drop. The angles of minimum deviations for red and violet colours are  $231^{\circ}$  and  $234^{\circ}$ , respectively, so they subtend a cone of  $51^{\circ}$  for the red and  $54^{\circ}$  for the violet colour. From Fig. 16.19 it is clear that the red colour will be on the inner and the violet colour on the outer side of the bow.

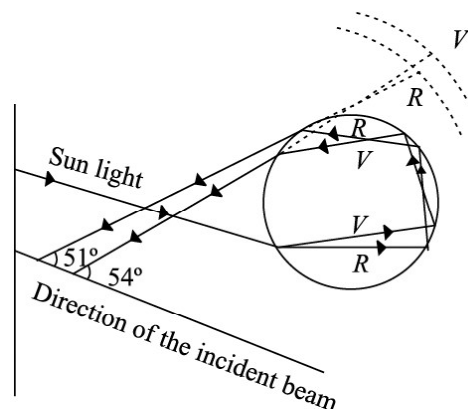


Fig. 16.19 : Formation of the secondary rainbow

The simultaneous appearance of the primary and secondary rainbows is shown in Fig. 16.20. The space between the two bows is relatively dark. Note that the secondary rainbow lies above the primary bow.

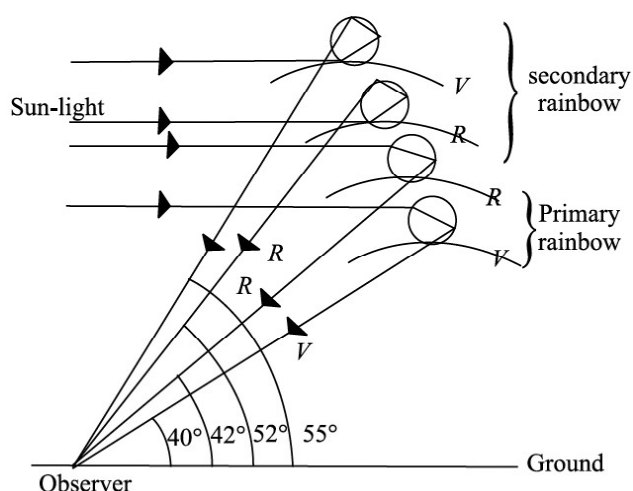


Fig. 16.20 : Simultaneous formation of the primary and secondary rainbow.

## 16.6) SCATTERING OF LIGHT IN ATMOSPHERE

On a clear day when we look at the sky, it appears blue. But the clouds appear white. Similarly, production of brilliant colours when sunlight passes through jewels and crystals also attracts our attention. You may like to know : How and why does it happen? These phenomena can be explained in terms of scattering of light. A solution of dust or particle-free benzene exposed to sunlight gives brilliant blue colour when looked sideways.

### 16.6.1 Scattering of Light

This phenomenon involves interaction of radiation with matter. Tiny dust particles are present in Earth's atmosphere. When sunlight falls on them, it gets diffused in all directions. That is why light reaches even those nooks and corners where it normally is not able to reach straight from the source.

Let us perform a simple activity.

#### Activity 16.1

Take a glass jar or a trough, fill it with water and add a little milk to it. Now allow a narrow beam of light from a white bulb to fall on it. Observe the light at  $90^\circ$ . You will see a bluish beam through water. This experiment shows that after scattering, the wavelengths of light become a peculiarly different in a given direction Fig. 16.21.

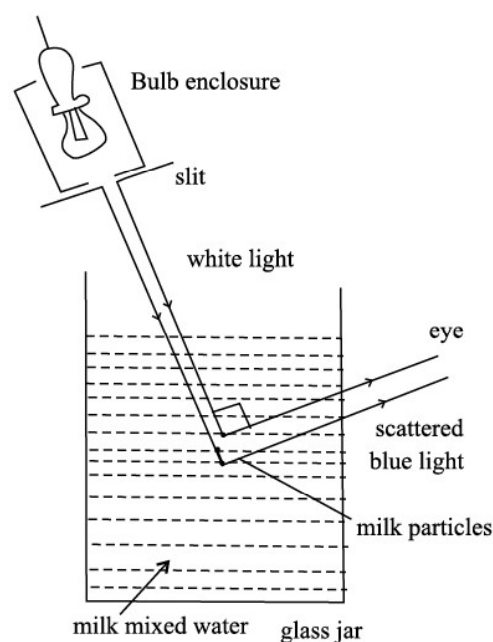


Fig. 16.21 : The scattering of light from milk particles

The phenomenon of scattering is a two step process : absorption of light by the scattering dispersion and scattering of particle and then instant re-emission by it in all possible directions. Thus, this phenomenon is different from reflection. The scattered light does not obey the laws of reflection. It is important to note that the size of the particle must be less than the wavelength of light incident on it. A bigger sized particle will scatter all the wavelengths equally. The intensity of scattered light is given by *Rayleigh's law* of scattering. According to this law, ***the intensity of scattered light is inversely proportional to the fourth power of its wavelength:***

$$I \propto \frac{1}{\lambda^4}$$

Here I is intensity and  $\lambda$  is wavelength of the scattered light. Thus, when white light is incident on the scattering particle, the blue light is scattered the most and the red light is scattered the least.

#### Example 16.4

Waves of wavelength 3934Å, 5890Å and 6867Å are found in the scattered beam when sunlight is incident on a thin layer of chimney smoke. Which of these is scattered more intensely?

#### Solution :

The intensity of scattered light is given by

$$I \propto \frac{1}{\lambda^4}$$

Since 3934Å is the smallest wavelength, it will be scattered most intensely. On the basis of scattering of light, we can explain why sky appears blue, clouds appear white and the sun appears red at sunrise as well as at sunset.

### C.V. Raman (1888 – 1970)

Chandra Shekhar Venkat Raman is the only Indian national to receive Nobel prize (1930) in Physics till date. His love for Physics was so intense that he resigned his job of an officer in Indian finance department and accepted the post of Palit Professor of Physics at the Department of Physics, Calcutta University. His main contributions are : Raman effect on scattering of light, molecular diffraction of light, mechanical theory of bowed strings, diffraction of X-rays, theory of musical instruments and Physics of crystals.



As a Director of Indian Institute of Science, Bangalore and later as the founder Director of Raman Research Institute, he did yeoman's work for Indian science and put it on firm footing in the pre-independence period.

- (A) **Blue Colour of the Sky :** We know that scattering of light by air molecules, water droplets or dust particles present in the atmosphere can be explained in accordance with Rayleigh's law. The shorter wavelengths are scattered more than the longer wavelengths. Thus, the blue light is scattered almost six times more intensely than the red light as the wavelength of the blue light is roughly 0.7 times that of the red. The scattered light becomes rich in the shorter wavelengths of violet, blue and green colours. On further scattering, the violet light does not reach observer's eye as the eye is comparatively less sensitive to violet than blue and other wavelengths in its neighbourhood. So, when we look at the sky far away from the sun, it appears blue.

### Example 16.5

What will be the colour of the sky for an astronaut in a spaceship flying at a high altitude.

#### Solution :

At a high altitude, in the absence of dust particle and air molecules, the sunlight is not scattered. So, the sky will appear black.

- (B) **White colour of the clouds :** The clouds are formed by the assembly of small water drops whose size becomes more than the average wavelength of the visible light ( $5000\text{\AA}$ ). These droplets scatter all the wavelengths with almost equal intensity. The resultant scattered light is therefore white. So, a thin layer of clouds appears white. What about dense clouds?
- (C) **Red colour of the sun at sunrise and sunset :** We are now able to understand the red colour of the sun at sunrise and sunset. In the morning and evening when the sun is near the horizon, light has to travel a greater distance through the atmosphere. The violet and blue wavelengths are scattered by dust particles and air molecules at an angle of about  $90^\circ$ . The sunlight thus becomes devoid of shorter wavelengths and the longer wavelength of red colour reaches the observer Fig. 16.22. So the sun appears to us as red.

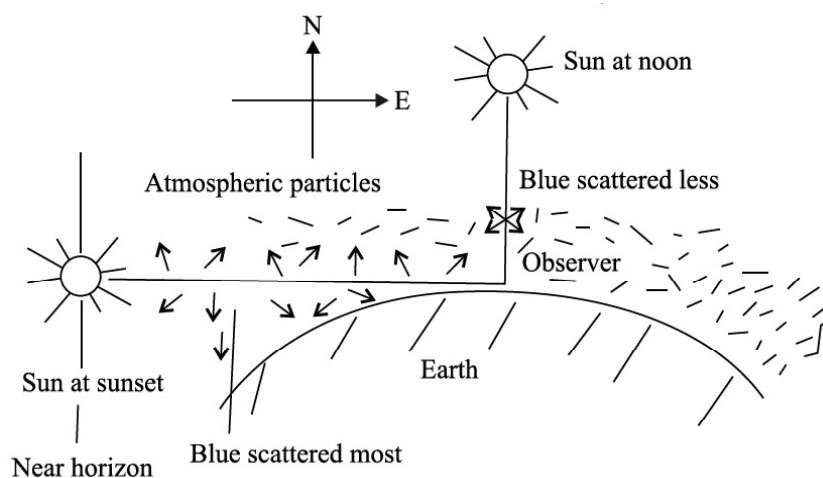


Fig. 16.22 : Red colour of the sun at sunset and sunrise (blue is scattered away)

At noon, the Sun is overhead and its distance from the observer is comparatively less. The blue colour is also scattered less. This results in the Sun appearing white, as a matter of fact, crimson.

### Intext Questions 16.6

1. Why dense clouds appear black?
2. Why does the sky appear deep blue after rains on a clear day?
3. Can you suggest an experiment to demonstrate the red colour of the Sun at sunrise and sunset?
4. The photographs taken from a satellite show the sky dark. Why?

### WHAT YOU HAVE LEARNT

- According to the Huygens' wave theory, light propagates in the form of wavefronts.
- The locus of all particles of the medium vibrating in the same phase at any instant of time is called the wavefront
- If two light sources emit light waves of the same frequency, same amplitude and move along the same path maintaining a constant phase difference between them, they are said to be coherent.
- When waves from two coherent sources superpose, a redistribution of energy takes place at different points. This is called the interference of light.
- For constructive interference, phase difference  $\delta = 2n\pi$  and for destructive interference, phase difference  $\delta = (2n + 1)\pi$ .
- The bending of light near the corners of an obstacle or aperture is called diffraction of light.
- The phenomenon in which vibrations of light get confined in a particular plane containing the direction of propagation is called polarisation of light.
- Light of single wavelength or colour is said to be monochromatic but sunlight, which has several colours or wavelengths, is polychromatic.
- The splitting of light into its constituent wavelengths on entering an optically denser medium is called dispersion.
- A prism is used to produce dispersed light, which when taken on the screen, forms the spectrum.
- The angle of deviation is minimum if the angles of incidence and emergence become equal. In this situation, the beam is most intense for that colour.
- The angle of deviation and refractive index for a small-angled prism are connected by the relation  $\delta = (\mu - 1)A$

- The rainbow is formed by dispersion of sunlight by raindrops at definite angles for each colour so that the condition of minimum deviation is satisfied.
- Rainbows are of two types : primary and secondary. The outer side of the primary rainbow is red but the inner side is violet. The remaining colours lie in between to follow the order (VIBGYOR). The scheme of colours gets reversed in the secondary rainbow.
- The blue colour of the sky, the white colour of clouds and the reddish colour of the sun at sunrise and sunset are due to scattering of light. The intensity of scattered light is inversely proportional to the fourth power of the wavelength  $\left( I \propto \frac{1}{\lambda^4} \right)$ . This is called Rayleigh's law. So the blue colour is scattered more than red.

### TERMINAL EXERCISE

1. Explain in brief the theories describing the nature of light.
2. What is a wavefront? What is the direction of a beam of light with respect to the associated wavefront? State the Huygens' principle and explain the propagation of light waves.
3. Obtain the laws of reflection on the basis of Huygens' wave theory.
4. What is the principle of superposition of waves? Explain the interference of light.
5. Describe Young's double slit experiment to produce interference. Deduce an expression for the width of the interference fringes.
6. What would happen to the interference pattern obtained in the Young's double slit experiment when
  - (i) one of the slits is closed
  - (ii) the experiment is performed in water instead of air
  - (iii) the source of yellow light is used in place of the green light source
  - (iv) the separation between the two slits is gradually increased
  - (v) white light is used in place of a monochromatic light
  - (vi) the separation between the slits and the screen is increased
  - (vii) two slits are slightly moved closer
  - (viii) each slit width is increased.
7. In Young's experimental set-up, the slit separation is 2 mm and the distance between the slits and the observation screen is 100 cm. Calculate the path difference between the waves arriving at a point 5 cm away from the point where the line dividing the slits touches the screen.
8. With the help of Huygens' construction, explain the phenomenon of diffraction.

9. How would you demonstrate that the light waves are transverse in nature?
10. Distinguish between the polarized and unpolarized lights.
11. State and explain Brewster's law.
12. The polarising angle for a medium is  $60^\circ$ . Calculate the refractive index.
13. For a material of refractive index 1.42, calculate the polarising angle for a beam of unpolarised light incident on it.
14. For a prism, show that  $i + e = A + \delta$ .
15. Would you prefer small-angled or a large-angled prism to produce dispersion. Why?
16. Under what condition is the deviation caused by a prism directly proportional to its refractive index?
17. Explain why the sea water appears blue at high seas.
18. The angle of minimum deviation for a  $60^\circ$  glass prism is  $39^\circ$ . Calculate the refractive index of glass.
19. The deviation produced for red, yellow and violet colours by a crown glass are  $2.84^\circ$ ,  $3.28^\circ$  and  $3.72^\circ$  respectively. Calculate the dispersive power of the glass material.
20. Calculate the dispersive power for flint glass for the following data :  $\mu_C = 1.6444$ ,  $\mu_D = 1.6520$ ,  $\mu_F = 1.6637$ , where C, D & F are the Fraunhofer nomenclatures.
21. A lens can be viewed as a combination of two prisms placed with their bases together. Can we observe dispersion using a lens. Justify your answer.
22. Human eye has a convex lens. Do we observe dispersion with unaided eye?

### ANSWERS TO INTEXT QUESTIONS

#### 16.1

1. Perpendicular to each other ( $\theta = \pi/2$ )
2.  $\frac{1}{2}$

#### 16.2

1. On the amplitude of the waves and the phase difference between them.
2. When the phase difference between the two superposing beams is an integral multiple of  $2\pi$ , we obtain constructive interference.
3. No, it is so because two independent sources of light will emit light waves with different wavelengths, amplitudes and the two set of waves will not have constant phase relationship. Such sources of light are called incoherent sources. For observing interference of light, the sources of light must be coherent. When the light waves are coming from two incoherent sources, the points on the screen where two crests



or two trough superpose at one instant to produce brightness may receive, at the other instant, the crest of the wave from one source and trough from the other and produce darkness. Thus, the whole screen will appear uniformly illuminated if the pinholes  $S_1$  and  $S_2$  are replaced by two incandescent light bulbs.

4. Coherent sources should emit waves
  - (a) of same frequency and wavelength
  - (b) in phase or having constant phase difference
  - (c) same amplitude and period.

No, our eyes cannot act as coherent sources of light. Coherent sources typically refer to light sources that emit waves with a consistent phase relationship, such as lasers or certain types of light sources. The human eye is not designed to emit coherent light waves; rather, it is an organ designed for receiving and detecting light.

### 16.3

1. Yes
2. Interference is the superposition of secondary waves emanating from two different secondary sources whereas diffraction is the superposition of secondary waves emanating from different portions of the same wavefronts.
3. Due to the increasing path difference between wavelets.

### 16.4

1. Yes. Because, in a longitudinal wave, the direction of vibrations is the same as the direction of motion of the wave.
2. No.
3.  $90^\circ$  or  $270^\circ$
4. No.

### 16.5

1. The velocity of propagation of waves of different wavelengths of visible light is almost the same in most ordinary gases. Hence, they do not disperse visible light. Their refractive index is also very close to 1.
2. Violet, because  $\lambda_r > \lambda_v$  and the velocity of the red light is more than that of the violet light inside an optically denser medium.
3. Yes

$$4. \quad \mu = \frac{\sin 60^\circ}{\sin 30^\circ} = \sqrt{3} = 1.732$$

### 16.6

1. It absorbs sunlight
2. It becomes clear of dust particles and bigger water molecules. The scattering now takes place strictly according to Rayleigh's law.
3. We can take sodium thiosulphate solution in a round bottom flask and add a small quantity of sulphuric acid. On illuminating this solution with a high power bulb, we can see a scenario similar to the colour of the sun at sunrise and sunset.
4. At very high altitudes no centres (particles) of scattering of sunlight are present. So the sky appears dark.

### ANSWERS TO TERMINAL EXERCISE

7. 0.1 mm
12. 1.73
13.  $54^\circ$
18. 1.5
19. 0.27
20. 0.03



# ELECTRICAL CHARGES AND FIELDS

## INTRODUCTION

So far you have learnt about mechanical, thermal and optical systems and various phenomena exhibited by them. The importance of electricity in our daily life is too evident. The physical comforts we enjoy and the various devices used in daily life depend on the availability of electrical energy. An electrical power failure demonstrates directly our dependence on electric and magnetic phenomena; the lights go off, the fans, coolers and air-conditioners in summer and heaters and gysers in winter stop working. Similarly, radio, TV, computers, microwaves can not be operated. Water pumps stop running and fields cannot be irrigated. Even train services are affected by power failure. Machines in industrial units cannot be operated. In short, life almost comes to a stand still, sometimes even evoking public anger. It is, therefore, extremely important to study electric and magnetic phenomena.

In this lesson, you will learn about two kinds of electric charges, their behaviour in different circumstances, the forces that act between them, the behaviour of the surrounding space etc. Broadly speaking, we wish to study that branch of physics which deals with electrical charges at rest. This branch is called **electrostatics**.

### OBJECTIVES

After studying this lesson, you should be able to

- *state the basic properties of electric charges;*
- *explain the concepts of quantisation and conservation of charge;*
- *explain Coulomb's law of force between electric charges;*
- *define electric field due to a point charge at rest and draw electric lines of force;*
- *define electric dipole, dipole moment and the electric field due to a dipole; and*
- *state Gauss' theorem and derive expressions for the electric field due to a point charge and a long charged wire.*

### 17.1 FRICTIONAL ELECTRICITY

The ancient Greeks observed electric and magnetic phenomena as early as 600 B.C. They found that a piece of amber, when rubbed, becomes electrified and attracts small pieces of feathers. The word **electric** comes from Greek word for amber meaning **electron**.

You can perform simple activities to demonstrate the existence of charges and forces between them. If you run a comb through your dry hair, you will note that the comb begins to attract small pieces of paper. Do you know how does it happen? Let us perform two simple experiments to understand the reason.

**Activity 17.1**

Take a hard rubber rod and rub it with fur or wool. Next you take a glass rod and rub it with silk. Suspend them (rubber rod and a glass rod) separately with the help of non-metallic threads, as shown in Fig. 17.1 and Fig. 17.2.

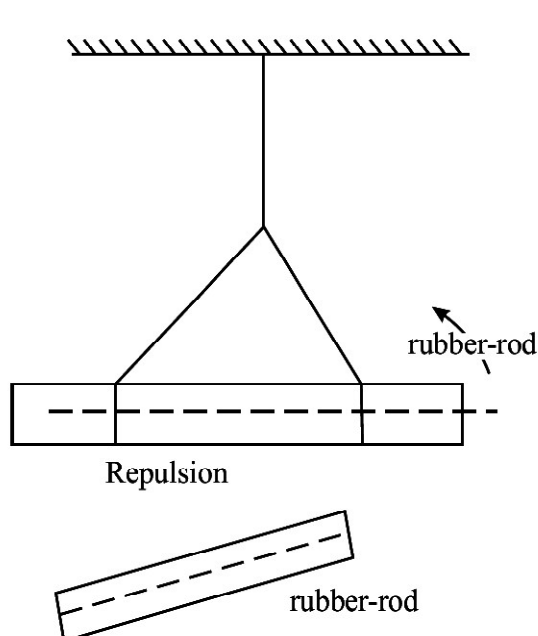


Fig. 17.1 : Force of attraction

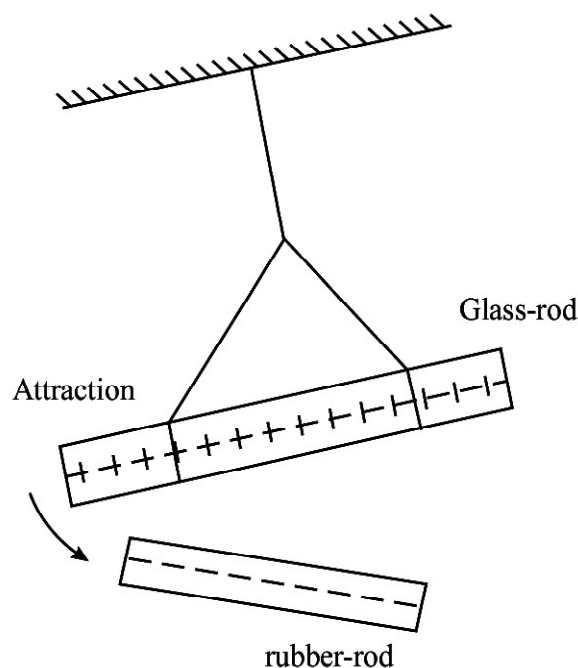


Fig. 17.2 : Force of repulsion

17.1 a charged rubber rod repels another charged rubber rod : like charges repel each other; and  
17.2 a charged glass rod attracts a charged rubber rod : unlike charges attract each other.

Now bring rubber rod rubbed with wool near these rods one by one. What do you observe? You will observe that

- When a charged rubber rod is brought near the charged (suspended) rubber rod, they show repulsion (Fig. 17.1); and
- When the charged rubber rod is brought near the (suspended) charged glass rod, they show attraction (Fig. 17.2).

Similar results will be obtained by bringing a charged glass rod.

On the basis of these observations, we can say that

- A charged rubber rod attracts a charged glass rod but repels a charged rubber rod.
- A charged glass rod repels a charged glass rod but attracts a charged rubber rod.

From these activities we can infer that the rubber rod has acquired one kind of electricity and the glass rod has acquired another kind of electricity. Moreover, **like charges repel and unlike charges attract each other.**

Franklin (Benjamin Franklin, 1706-1790) suggested that the charge on the glass rod is to be called **positive** and that on the rubber rod is to be called **negative**. We follow this convention since then.

Once a body is charged by friction, it can be used to charge other conducting bodies by *conduction*, i.e., by touching the charged body with an uncharged body; and *induction*, i.e., by bringing the charged body close to an uncharged conductor and earthing it. Subsequently, the charged body and the earthing are removed simultaneously.

### Conservation of Charge

In Activity 17.1, you have seen that when a glass rod is rubbed with silk, the rod acquires positive charge and silk acquires negative charge. Since both materials in the normal state are neutral (no charge), the positive charge on the glass rod should be equal in magnitude to the negative charge on silk. This means that the total charge of the system (glass + silk) is conserved. Charge is neither created nor destroyed. It is only transferred from one body of the system to the other. The transfer of charges takes place due to increase in the thermal energy of the system when the glass rod is rubbed; the less tightly bound electrons from the glass rod are transferred to silk. The glass rod (deficient in electrons) becomes positively charged and silk, which now has excess electrons, becomes negatively charged. When rubber is rubbed with fur, electrons from the fur are transferred to rubber. That is, rubber gains negative charge and fur gains an equal amount of positive charge. Any other kind of charge (other than positive and negative) has not been found till today.

### Quantisation of Charge

In 1909, Millikan (Robert Millikan, 1886-1953) experimentally proved that charge always occurs as some integral multiple of a fundamental unit of charge, which is taken as the charge on an electron. This means that if  $Q$  is the charge on an object, it can be written as  $Q = Ne$ , where  $N$  is an integer and  $e$  is charge on an electron. Then we say that charge is quantised. It means that a charged body cannot have  $2.5 e$  or  $6.4 e$  amount of charge. In Nuclear Physics you will learn that an electron has charge  $-e$  and a proton has charge  $+e$ . Neutron has no charge. Every atom has equal number of electrons and protons and that is why it is neutral. From this discussion, we can draw the following conclusions :

- There are only two kinds of charges in nature; positive and negative.
- Charge is conserved. Charge neither be created nor destroyed, but charge can be transferred from one body to another.
- Charge is quantised. Charge is not continuous in nature. It can be only certain amounts of charge is possible based on

$$Q = Ne \text{ like}$$

$$Q_1 = 1 \times 1.6 \times 10^{-19} = 1.6 \times 10^{-19} \text{ C}$$

$$Q_2 = 2 \times 1.6 \times 10^{-19} = 3.2 \times 10^{-19} \text{ C}$$

$Q_3 = 3 \times 1.6 \times 10^{-19} = 4.8 \times 10^{-19} \text{ C}$  and so on. Intermediate values of charges like  $2 \times 10^{-19} \text{ C}$ , or  $3 \times 10^{-19} \text{ C}$  or  $4 \times 10^{-19} \text{ C}$  so on are not possible.

**Intext Questions 17.1**

1. A glass rod when rubbed with silk cloth acquires a charge  $q = +3.2 \times 10^{-17}$  C.
  - i) Is silk cloth also charged?
  - ii) What is the nature and magnitude of the charge on silk cloth?
2. There are two identical metallic spheres A and B. A is given a charge  $+Q$ . Both spheres are then brought in contact and then separated.
  - (i) Will there be any charge on B ?
  - (ii) What will the magnitude of charge on B, if it gets charged when in contact with A.
3. A charged object has  $q = 4.8 \times 10^{-16}$  C. How many units of fundamental charge are there on the object? (Take  $e = 1.6 \times 10^{-19}$  C).

**17.2 COULOMB'S LAW**

You have learnt that two stationary charges either attract or repel each other. The force of attraction or repulsion between them depends on their nature. Coulomb studied the nature of this force and in 1785 established a fundamental law governing it. From experimental observations, he showed that the electrical force between two static point charges  $q_1$  and  $q_2$  placed some distance  $r$  apart is

- directly proportional to the product of their charges  $q_1, q_2$  ;
- inversely proportional to the square of the distance  $r$  between them ;
- directed along the line joining the two charged particles ; and
- repulsive for same kind of charges and attractive for opposite charges.

The magnitude of force  $F$  can then be expressed as

$$F = k \frac{q_1 \times q_2}{r^2} \quad (17.1)$$

For free space, we write 
$$F = \frac{1}{4\pi\epsilon_0} \times \frac{q_1 \times q_2}{r^2} \quad (17.2)$$

where constant of proportionality  $k = \frac{1}{4\pi\epsilon_0}$  for free space (vacuum) and  $k = \frac{1}{4\pi\epsilon}$  for a

material medium.  $\epsilon_0$  is called permittivity of free space and  $\epsilon$  is the permittivity of the medium. It means that for the same system of charges kept in a material medium, the magnitude of Coulomb force will be different from that placed in free space.

The constant  $k$  has a value which depends on the units of the quantities involved. The unit of charge in SI system is coulomb (C). The coulomb is defined in terms of the unit of current, called **ampere**. (You will learn about it later.) In SI system of units, the value of  $k$  is

$$k = \frac{1}{4\pi\epsilon_0} = 9 \times 10^9 \frac{\text{Nm}^2}{\text{C}^2} \quad (17.3)$$

since  $\epsilon_0 = 8.85 \times 10^{-12} \text{ C}^2 \text{ N}^{-1} \text{ m}^{-2}$ .

Thus in terms of force, one coulomb charge can be defined as : ***If two equal charges separated by one metre experience a force of  $9 \times 10^9 \text{ N}$ , then each charge has a magnitude of one coulomb.*** The value of electronic charge  $e$  is  $1.60 \times 10^{-19} \text{ C}$ . Note that

- Coulomb's law is also an inverse square law just like Newton's law of Gravitation,
- Coulomb's law holds good for point charges only.
- Coulomb's force acts at a distance, unlike mechanical force.

### How Big is One Coulomb?

The unit of electrical charge is coulomb. Have you ever thought : How big a coulomb is? To know this, let us calculate the magnitude of force between two charges, each of one coulomb, placed at a distance of one metre from one another:

$$\begin{aligned} |F| &= k \times \frac{q_1 \times q_2}{r^2} \\ &= 9.0 \times 10^9 \times \frac{1 \times 1}{1^2} \\ &= 9.0 \times 10^9 \approx 10^{10} \text{ N} \end{aligned}$$

If the mass of a loaded passenger bus is 5000 kg,  
its weight  $mg = (5000 \times 10) \text{ N}$  (assume  $g \approx 10 \text{ m s}^{-2}$ ) =  $5 \times 10^4 \text{ N}$ .

Let us assume that there are 10,000 such loaded buses in Delhi. The total weight of all these buses will be  $5 \times 10^4 \times 10,000 = 5 \times 10^8 \text{ N}$ . If there are 10 cities having same number of buses as those in Delhi, the total weight of all these loaded buses will be  $5 \times 10^9 \text{ N}$ . It means that the force between two charges, each of 1C and separated by one metre is equivalent to the weight of about two hundred thousand buses, each of mass 5000 kg.

### How many number of electrons make one Coulomb of charge?

$$\begin{aligned} Q &= 1 \text{ C} \\ n &= ? \\ e &= 1.6 \times 10^{-19} \text{ C} \\ Q &= ne \\ 1\text{C} &= n \times 1.6 \times 10^{-19} \end{aligned}$$

$$n = \frac{1}{1.6 \times 10^{-19}} = \frac{1 \times 10^{19}}{1.6}$$

$$n = \frac{10}{1.6} \times 10^{18}$$

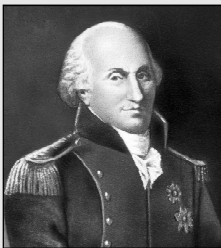
$$n = \frac{100}{16} \times 10^{18}$$

$$n = 6.25 \times 10^{18}$$

$$n = 625 \times 10^{16} \text{ electrons}$$

**Is Coulomb a bigger charge or charge of an electron?**

### Charles Augustin de Coulomb (1736– 1806)



A French physicist, Coulomb started his career as military engineer in West Indies. He invented a torsional balance and used it to perform experiments to determine the nature of interaction forces between charges and magnets. He presented the results of these experiments in the form of Coulomb's law of electrostatics and Coulomb's law of magnetostatics. The SI unit of charge has been named in his honour.

You now know that the ratio of forces between two point charges  $q_1$  and  $q_2$  separated by a distance  $r$ , when kept in free space (vacuum) and material medium, is equal to  $\varepsilon/\varepsilon_0$ :

$$\frac{F_0 \text{ (in vacuum)}}{F \text{ (in medium)}} = \frac{\varepsilon}{\varepsilon_0} = \varepsilon_r$$

where  $\varepsilon_r$  is known as relative permittivity or **dielectric constant**. Its value is always greater than one. We will define dielectric constant in another form later.

### Vector Form of Coulomb's Law

You know that force is a vector quantity. It means that force between two charges should also be represented as a vector. That is, Eqn. (17.1) should be expressed in vector form.

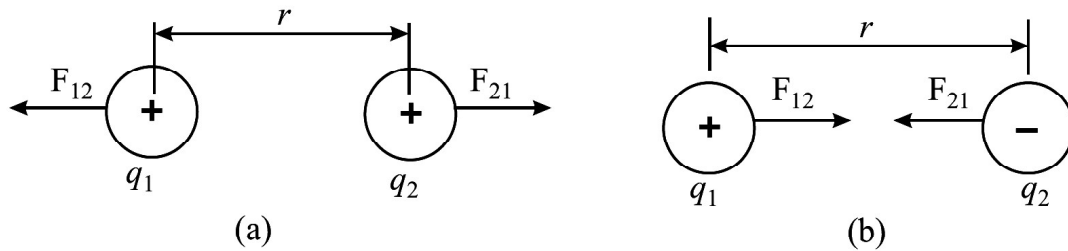
Let there be two point charges  $q_1$  and  $q_2$  separated by a distance  $r$  (Fig. 17.3). Suppose that  $\mathbf{F}_{12}$  denotes the force experienced by  $q_1$  due to the charge  $q_2$  and  $\mathbf{F}_{21}$  denotes the force on  $q_2$  due to charge  $q_1$ . We denote the unit vector pointing from  $q_1$  to  $q_2$  by  $\hat{\mathbf{r}}_{12}$ . Then from Fig. 17.3 (a), it follows that

$$\mathbf{F}_{12} = k \frac{q_1 q_2}{r_{12}^2} \hat{\mathbf{r}}_{12} \quad (17.4)$$

Similarly, for charges shown in Fig.17.3 (b), we can write



$$\mathbf{F}_{21} = -k \frac{q_1 q_2}{|r_{12}|^2} \hat{\mathbf{r}}_{12} \quad (17.5)$$



**Fig. 17.3 :** Two point charges  $q_1$  and  $q_2$  separated by a distance  $r$  :  
 (a) the direction of forces of repulsion between two positive charges, and  
 (b) the direction of forces of attraction between a positive and a negative charge.

The positive sign in Eqn. (17.4) indicates that the force is repulsive and the negative sign in Eqn. (17.5) indicates that the force is attractive (-ve sign for attractive force).

The Coulomb's law obeys the principle of action and reaction between two charges  $q_1$  and  $q_2$ . Therefore,

$$\mathbf{F}_{12} = -\mathbf{F}_{21} \quad (17.6)$$

In general, we can write the expression for force between two charges as

$$\mathbf{F}_{12} = k \times \frac{q_1 q_2}{r^2} \hat{\mathbf{r}}_{12} \quad (17.7)$$

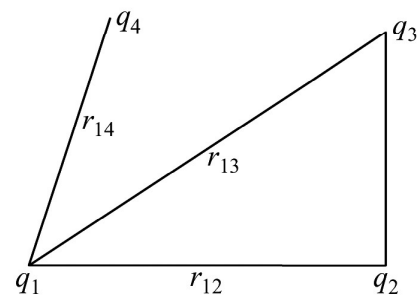
### Principle of Superposition

If there are more than two charges, we can calculate the force between any two charges using Eqn. (17.7). Suppose now that there are several charges  $q_1, q_2, q_3, q_4$ , etc. The force exerted on  $q_1$  due to all other charges is given by Eqn. (17.7):

$$\mathbf{F}_{12} = k \frac{q_1 q_2}{|r_{12}|^2} \hat{\mathbf{r}}_{12}$$

$$\mathbf{F}_{13} = k \frac{q_1 q_3}{|r_{13}|^2} \hat{\mathbf{r}}_{13}$$

and 
$$\mathbf{F}_{14} = k \frac{q_1 q_4}{|r_{14}|^2} \hat{\mathbf{r}}_{14} \quad (17.8)$$



**Fig. 17.4:** Principle of superposition

The resultant of all these forces, i.e., the total force  $\mathbf{F}$  experienced by  $q_1$  is their vector sum:

$$\mathbf{F} = \mathbf{F}_{12} + \mathbf{F}_{13} + \mathbf{F}_{14} + \dots \quad (17.9)$$

This is known as principle of superposition.

**Example 17.1**

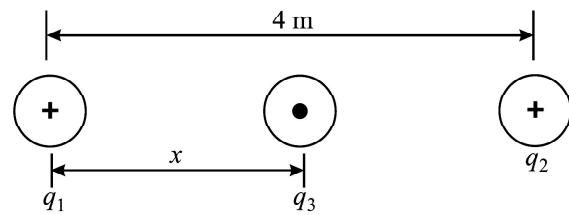
A charge  $+q_1 = 12\text{C}$  is placed at a distance of 4.0 m from another charge  $+q_2 = 6\text{C}$ , as shown in the Fig. 17.5. Where should a negative charge  $q_3$  be placed on the line joining  $q_1$  and  $q_2$  so that the charge  $q_3$  does not experience any force?

**Solution :**

Let  $q_3$  be placed between  $q_1$  and  $q_2$  at a distance of  $x$  metre from  $q_1$ . (It can be easily seen that on placing  $q_3$  on the left of  $q_1$  or on the right of  $q_2$  or at any position other than the one between the line joining  $q_1$  and  $q_2$ , the resultant force can not be zero.) The force exerted on  $q_3$  by  $q_1$  will be

$$\mathbf{F}_{31} = k \frac{q_1 q_3}{r_{31}^2} \hat{\mathbf{r}}_{31} \text{ towards } q_1$$

$$\therefore |\mathbf{F}_{31}| = k \frac{q_3 q_1}{x^2}$$



**Fig. 17.5 :** Three point charges  $q_1$ ,  $q_2$  and  $q_3$  placed in a straight line

The magnitude of force on  $q_3$  due to  $q_2$  is given by

$$|\mathbf{F}_{32}| = k \frac{q_3 q_2}{(4-x)^2} \text{ towards } q_2$$

The resultant force on  $q_3$  will be zero when  $\mathbf{F}_{31} = \mathbf{F}_{32}$ . Therefore, on substituting the numerical values, we get

$$k \times \frac{12q_3}{x^2} = k \times \frac{6q_3}{(4-x)^2}$$

Note that  $6q_3k$  is common on both sides and cancels out. Therefore, on simplification, we get

$$\frac{2}{x^2} = \frac{1}{(4-x)^2}$$

$$\begin{aligned} \text{or} \quad & 2(4-x)^2 = x^2 \\ \Rightarrow & x^2 - 16x + 32 = 0 \end{aligned}$$

On solving this, we get two values of  $x$  : 2.35 m and 13.65 m. The latter value is inadmissible because it goes beyond  $q_2$ . Therefore, the charge  $q_3$  should be placed at a distance of 2.35 m from  $q_1$ .

It is a reasonable solution qualitatively also. The charge  $q_1$  is stronger than  $q_2$ . Hence the distance between  $q_1$  and  $q_3$  should be greater than that between  $q_2$  and  $q_3$ .

**Example 17.2**

Two charges, each of  $6.0 \times 10^{-10}\text{C}$ , are separated by a distance of 2.0 m. Calculate the magnitude of Coulomb force between them.

**Solution :**

We know that the magnitude of Coulomb force between two charges is given by Eqn. (17.2) :

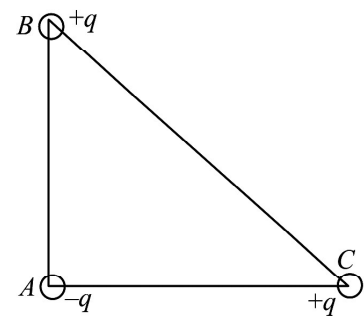
$$F = k \frac{q_1 \cdot q_2}{r^2}$$

Given,  $q_1 = q_2 = 6.0 \times 10^{-10} \text{C}$  and  $r = 2.0 \text{ m}$ , Therefore on putting these values, we get

$$\begin{aligned} F &= \frac{(9 \times 10^9 \text{ N m}^2 \text{C}^{-2}) \times (6.0 \times 10^{-10} \text{C})^2}{2^2 \text{ m}^2} \\ &= \frac{9 \times 10^9 \times 36.0 \times 10^{-20}}{4} \text{ N} \\ &= 81 \times 10^{-11} \text{ N} \end{aligned}$$

**Intext Questions 17.2**

- Two charges  $q_1 = 16 \mu\text{C}$  and  $q_2 = 9 \mu\text{C}$  are separated by a distance 12 m. Determine the magnitude of the force experienced by  $q_1$  due to  $q_2$  and also the direction of this force. What is the direction of the force experienced by  $q_2$  due to  $q_1$ ?
- There are three point charges of equal magnitude  $q$  placed at the three corners of a right angle triangle, as shown in  $AB = AC$ . What is the magnitude and direction of the force exerted on  $-q$ ?



Three charges placed at the three corners of a right angle triangle.

**17.3) ELECTRIC FIELD**

To explain the interaction between two charges placed at a distance, Faraday introduced the concept of electric field. The electric field  $\mathbf{E}$  at a point is defined as the electric force  $\mathbf{F}$  experienced by a positive test charge  $q_0$  placed at that point divided by the magnitude of the test charge. Mathematically, we write

$$\mathbf{E} = \frac{\mathbf{F}}{q_0} \quad (17.10)$$

This is analogous to the definition of acceleration due to gravity,  $\mathbf{g} = \mathbf{F}/m_0$ , experienced by mass  $m_0$  in the gravitational field  $\mathbf{F}$ .

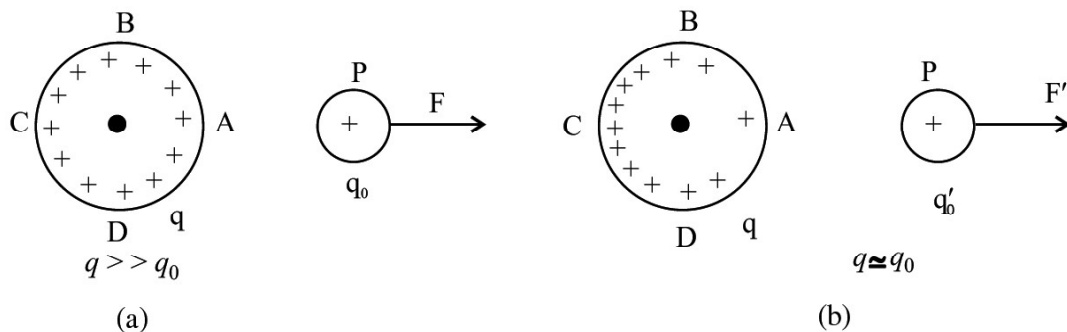
The electric field  $\mathbf{E}$  is a vector quantity and has the same direction as the electric force  $\mathbf{F}$ . Note that the electric field is due to an external charge and not due to the test charge. The test charge  $q_0$  should, therefore, be so small in magnitude that it does not disturb the

field due to external charge. (In practice, however, even the smallest test charge will disturb the external field.) Strictly speaking, mathematical definition given below is more accurate:

$$\mathbf{E} = \lim_{q_0 \rightarrow 0} \frac{\mathbf{F}}{q_0} \quad (17.11)$$

In SI system, the force is in newton and the charge is in coulomb. Therefore, according to Eqn. (17.10), the electric field has the unit newton per coulomb. The direction of  $\mathbf{E}$  is same as that of  $\mathbf{F}$ . Note that *the action of electric force is mediated through electric field*.

Let us now examine why the test charge  $q_0$  should be infinitesimally small.



**Fig. 17.6 :** (a) uniformly charged metallic sphere and a test charge, and (b) redistribution of charge on the sphere when another charge is brought near it.

Refer to Fig. 17.6. It shows a uniformly charged metallic sphere with charge  $q$  and a test charge  $q_0 (<< q)$ . It means that charge density per unit area is same around points A, B, C and D. The test charge  $q_0$  must measure the force  $\mathbf{F}$  without disturbing the charge distribution on the sphere. Fig. 17.6 (b) shows the situation when  $q \simeq q_0$ . In this case, the presence of the test charge modifies the surface charge density. As a result, the electrical force experienced by the test charge  $q_0$  will also change, say from  $\mathbf{F}$  to  $\mathbf{F}'$ . That is, the force in the presence of test charge is different from that in its absence. But without  $q_0$ , the force cannot be measured. If  $q_0$  is infinitesimally small in comparison to  $q$ , the charge distribution on the sphere will be minimally affected and the results of measurement will have a value very close to the true value. That is,  $\mathbf{F}'$  will be very nearly equal to  $\mathbf{F}$ . We hope you now appreciate the point as to why the test charge should be infinitesimally small.

Let there be a point charge  $q$ . A test charge  $q_0$  is placed at a distance  $r$  from  $q$ . The force experienced by the test charge is given by

$$\mathbf{F} = k \frac{qq_0}{r^2} \hat{\mathbf{r}} \quad (17.12)$$

The electric field is defined as the force per unit charge. Hence

$$\mathbf{E} = k \times \frac{q}{r^2} \hat{\mathbf{r}} \quad (17.13)$$

If  $q$  is positive, the field  $\mathbf{E}$  will be directed away from it. If  $q$  is negative, the field  $\mathbf{E}$  will be directed towards it. This is shown in Fig. 17.7.

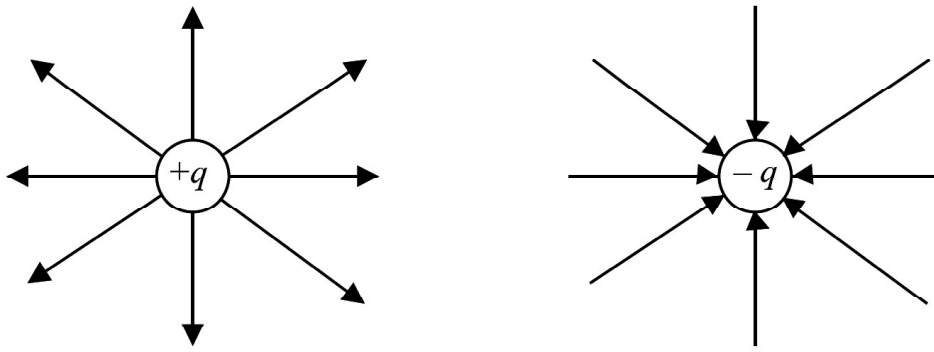


Fig. 17.7 : Direction of electric field due to positive and negative charges

The principle of superposition applies to electric field also. If there are a number of charges  $q_1, q_2, q_3, \dots$ , the corresponding fields at a point P according to Eqn. (17.13) are

$$\mathbf{E}_1 = k \times \frac{q_1}{r_1^2} \hat{\mathbf{r}}_1, \quad \mathbf{E}_2 = k \times \frac{q_2}{r_2^2} \hat{\mathbf{r}}_2 \quad \text{and} \quad \mathbf{E}_3 = k \times \frac{q_3}{r_3^2} \hat{\mathbf{r}}_3 \quad (17.14)$$

The total field at point P due to all charges is the vector sum of all fields. Thus,

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 + \dots$$

or

$$\mathbf{E} = k \sum_{i=1}^N \frac{q_i}{r_i^2} \hat{\mathbf{r}}_i \quad (17.15)$$

where  $r_i$  is the distance between P and charge  $q_i$  and  $\hat{\mathbf{r}}_i$  is the unit vector directed from  $q_i$  to P. The force on a charge  $q$  in an electric field  $\mathbf{E}$  is

$$\mathbf{F} = q\mathbf{E} \quad (17.16)$$

### Example 17.3

The electric force at some point due to a point charge  $q = 3.5\mu\text{C}$  is  $8.5 \times 10^{-4} \text{ N}$ . Calculate the strength of electric field at that point.

**Solution :**

From Eq. (17.16) we can write

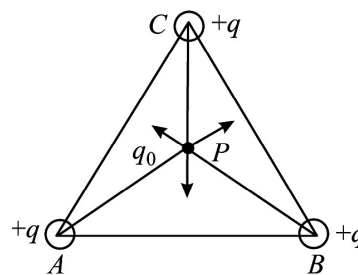
$$\begin{aligned} E &= \frac{F}{q} = \frac{8.5 \times 10^{-4} \text{ N}}{3.5 \times 10^{-6} \text{ C}} \\ &= 2.43 \times 10^2 \text{ NC}^{-1} \end{aligned}$$

### Example 17.4

Three equal positive point charges are placed at the three corners of an equilateral triangle, as shown in Fig. 17.8. Calculate the electric field at the centroid P of the triangle.

**Solution :**

Suppose that a test charge  $q_0$  has been placed at the centroid  $P$  of the triangle. The test charge will experience force in three directions making same angle between any two of them. The resultant of these forces at  $P$  will be zero. Hence the field at  $P$  is zero.

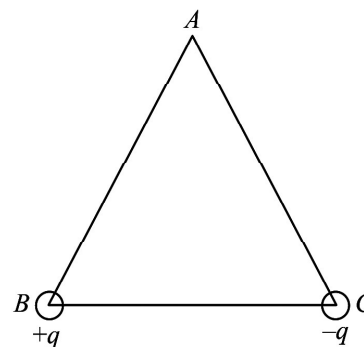


**Fig. 17.8 :** Electric field at the centroid of an equilateral triangle due to equal charges at its three corners is zero.

**Intext Questions 17.3**

1. A charge  $+Q$  is placed at the origin of co-ordinate system. Determine the direction of the field at a point  $P$  located on  
 (a)  $+x$ -axis (b)  $+y$ -axis (c)  $x = 4$  units and (d)  $y = 4$  units

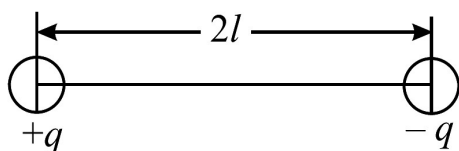
2. The  $\Delta ABC$  is defined by  $AB = AC = 40$  cm. And angle at  $A$  is  $30^\circ$ . Two charges, each of magnitude  $2 \times 10^{-6}$  C but opposite in sign, are placed at  $B$  and  $C$ , as shown in Fig. 17.9. Calculate the magnitude and direction of the field at  $A$ .



**Fig. 17.9**

3. A negative charge is located in space and the electric field is directed towards the earth. What is the direction of the force on this charge?
4. Two identical charges are placed on a plane surface separated by a distance  $d$  between them. Where will the resultant field be zero?

**Electric Field due to a Dipole**



**Fig. 17.10 :** Two unlike charges of equal magnitude separated by a small distance form a dipole.

If two equal and opposite charges are separated by a small distance, the system is said to form a dipole. The most familiar example is  $H_2O$ . Fig. 17.10 shows charges  $+q$  and  $-q$  separated by a small distance  $2l$ . The product of the magnitude of charge and separation between the charges is called **dipole moment,  $P$**  :

$$P = q \times 2l \tag{17.17}$$

Its SI unit is coulomb-metre.

The dipole moment is a vector quantity. Eqn. (17.17) gives its magnitude and its direction is from negative charge to positive charge along the line joining the two charges (axis of the dipole). Having defined a dipole and dipole moment, we are now in a position to calculate

the **electric field due to a dipole**. The calculations are particularly simple in the following cases.

**CASE I :** *Electric field due to a dipole at an axial point : End-on position*

To derive an expression for the electric field of a dipole at a point P which lies on the axis of the dipole, refer to Fig. 17.11. This is known as **end-on position**. The point charges  $-q$  and  $+q$  at points A and B are separated by a distance  $2l$ . The point O is at the middle of AB. Suppose that point P is at a distance  $r$  from the mid point O. Then electric field at P due to  $+q$  at B is given by

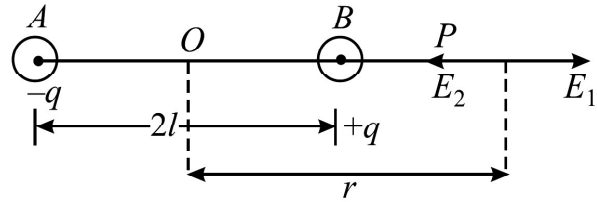


Fig. 17.11 : Field at point P on the dipole axis

$$\mathbf{E}_1 = k \times \frac{q}{(r-l)^2} \text{ in the direction AP}$$

Similarly, the electric field  $\mathbf{E}_2$  at P due to  $-q$  is given by

$$\mathbf{E}_2 = k \times \frac{q}{(r+l)^2} \text{ in the direction PA}$$

The resultant field  $\mathbf{E}$  at P will be in the direction of  $\mathbf{E}_1$ , since  $\mathbf{E}_1$  is greater than  $\mathbf{E}_2$  [as  $(r-l)$  is less than  $(r+l)$ ]. Hence

$$\begin{aligned} \mathbf{E} &= \frac{kq}{(r-l)^2} - \frac{kq}{(r+l)^2} \\ &= kq \left[ \frac{1}{(r-l)^2} - \frac{1}{(r+l)^2} \right] \\ &= kq \left[ \frac{(r+l)^2 - (r-l)^2}{(r^2 - l^2)^2} \right] \\ &= kq \times \frac{4lr}{(r^2 - l^2)^2} \\ &= k \frac{(2lq)2r}{(r^2 - l^2)^2} \\ &= k \frac{2Pr}{(r^2 - l^2)^2} \end{aligned}$$

where dipole moment  $\mathbf{P} = 2lq$ . Since  $k = 1/4\pi\epsilon_0$ , we can rewrite it as

$$\mathbf{E} = \frac{2\mathbf{P}}{4\pi\epsilon_0} \times \frac{r}{r^4(1-l^2/r^2)^2}$$

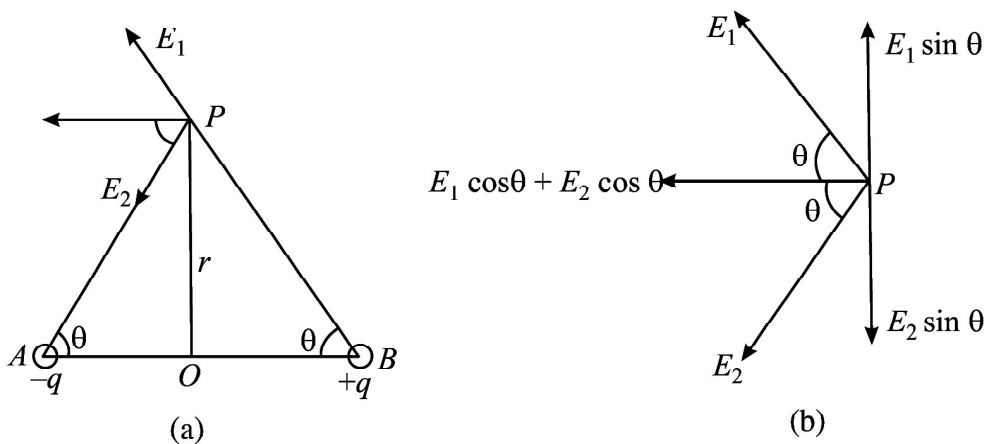
If  $r \gg l$ ,  $l^2/r^2$  will be very small compared to  $l$ . It can even be neglected and the expression for electric field then simplifies to

$$\mathbf{E} = \frac{2\mathbf{P}}{4\pi\epsilon_0 r^3} \quad (17.18)$$

It shows that electric field is in the direction of  $\mathbf{P}$  and its magnitude is inversely proportional to the third power of distance of the observation point from the centre of the dipole.

**CASE II :** *Electric field due to a dipole at a point on the perpendicular bisector : Broad-on position*

Suppose that point P lies on the perpendicular bisector of the line joining the charges shown in Fig. 17.12. Note that  $AB = 2l$ ,  $OP = r$ , and  $AO = OB = l$ .



**Fig. 17.12 :** (a) Field at point P on the perpendicular bisector of the line joining the charges, and  
(b) resolution of field in rectangular components.

The angle  $\theta$  is shown in Fig. 17.12 (a). From right  $\Delta$ s PAO and PBO, we can write

$$AP = BP = \sqrt{l^2 + r^2}$$

The field at P due to charge  $+q$  at B in the direction of BP can be written as

$$E_1 = k \frac{q}{l^2 + r^2}$$

Similarly, the field at P due to charge at A in the direction of PA is given as

$$E_2 = k \frac{q}{l^2 + r^2}$$



Note that the magnitudes of  $\mathbf{E}_1$  and  $\mathbf{E}_2$  are equal.

Let us resolve the fields  $\mathbf{E}_1$  and  $\mathbf{E}_2$  parallel and perpendicular to AB. The components parallel to AB are  $\mathbf{E}_1 \cos \theta$  and  $\mathbf{E}_2 \cos \theta$ , and both point in the same direction.

The components normal to AB are  $\mathbf{E}_1 \sin \theta$  and  $\mathbf{E}_2 \sin \theta$  and point in opposite directions. [Fig. 17.12 (b)]. Since these component are equal in magnitude but opposite in direction, they cancel each other. Hence, the magnitude of resultant electric field at P is given by

$$\begin{aligned} E &= E_1 \cos \theta + E_2 \cos \theta \\ &= k \frac{q}{l^2 + r^2} \cos \theta + k \frac{q}{l^2 + r^2} \cos \theta \end{aligned}$$

But  $\cos \theta = \frac{1}{\sqrt{(l^2 + r^2)}}$ . Using this expression in the above result, the electric field at

P is given by

$$\begin{aligned} E &= \frac{kq}{(l^2 + r^2)} \times \frac{2l}{\sqrt{(l^2 + r^2)}} \\ &= k \frac{2lq}{(l^2 + r^2)^{3/2}} \\ &= k \frac{2lq}{r^3 (1 + l^2/r^2)^{3/2}} \end{aligned}$$

But  $\mathbf{P} = 2lq$ . If  $r^2 \gg l^2$ , the factor  $l^2/r^2$  can be neglected in comparison to unity. Hence

$$E = \frac{\mathbf{P}}{4\pi\epsilon_0 r^3} \quad (17.19)$$

Note that electric field due to a dipole at a point in broad-on position is inversely proportional to the third power of the perpendicular distance between P and the line joining the charges. The direction of electrical field 'E' is exactly opposite to the direction of electric dipole moment.

If we compare Eqns. (17.18) and (17.19), we note that the electric field in both cases is proportional to  $1/r^3$ . But there are differences in details:

- The magnitude of electric field in end-on-position is twice the field in the broad-on position.
- The direction of the field in the end-on position is along the direction of dipole moment, whereas in the broad-on position, it is oppositely directed.

## Electric Dipole in a Uniform Field

A uniform electric field has constant magnitude and fixed direction. Such a field is produced between the plates of a charged parallel plate capacitor. Pictorially, it is represented by equidistant parallel lines. Let us now examine the behaviour of an electric dipole when it is placed in a uniform electric field (Fig. 17.13).

Let us choose x-axis such that the electric field points along it. Suppose that the dipole axis makes an angle  $\theta$  with the field direction. A force  $qE$  acts on charge  $+q$  along the  $+x$  direction and an equal force acts on charge  $-q$  in the  $-x$  direction. Two equal, unlike and parallel forces separated by a perpendicular distance  $y$  form a couple. For couple to form two equal unlike and parallel forces must not lie in the same line. They must be separated by perpendicular distance called as arm of couple. Then only they form a couple. (If two equal, unlike and parallel forces lie in the same line they cancel each other and net force is zero). The couple tends to align the dipole in the direction of the external electric field  $E$ . The magnitude of torque  $\tau$  is given by

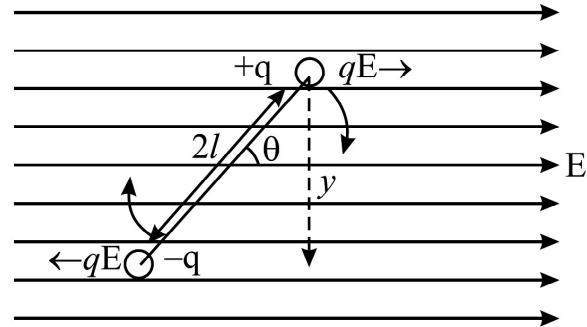


Fig. 17.13 : A dipole in a uniform electric field. The forces on the dipole form a couple and tend to rotate it.

The magnitude of torque  $\tau$  is given by

$$\begin{aligned}\tau &= \text{Force} \times \text{arm of the couple} \\ &= qE \times y \\ &= qE \times 2l \sin \theta \\ &= pE \sin \theta\end{aligned}$$

In vector form, we can express this result to

$$\tau = \mathbf{p} \times \mathbf{E} \quad (17.20)$$

We note that

- when  $\theta = 0$ , the torque is zero, and
- for  $\theta = 90^\circ$ , the torque on the dipole is maximum, equal to  $pE$ . So we may conclude that the electric field tends to rotate the dipole and align it along its own direction.

### Example 17.5

Two charges  $+q$  and  $-q$ , each of magnitude  $6.0 \times 10^{-6}$  C, form a dipole. The separation between the charges is  $4 \times 10^{-10}$  m. Calculate the dipole moment. If this dipole is placed in a uniform electric field  $E = 3.0 \times 10^2$  NC $^{-1}$  at an angle  $30^\circ$  with the field, calculate the value of torque on the dipole.

**Solution :**

$$\begin{aligned}\text{The dipole moment } p &= qd \\ &= (6.0 \times 10^{-6}\text{C}) \times (4.0 \times 10^{-10} \text{ m}) \\ &= 24 \times 10^{-16} \text{ Cm.}\end{aligned}$$

Since torque  $\tau = pE \sin \theta$ , we can write

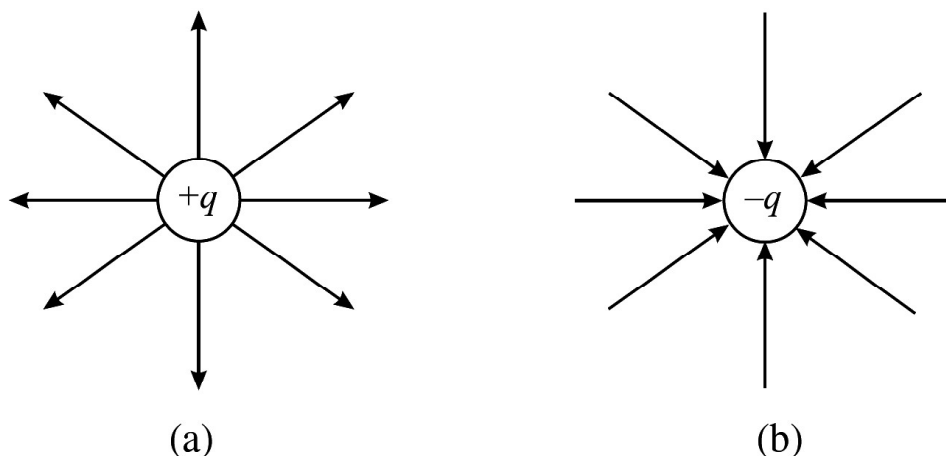
$$\begin{aligned}\tau &= (24 \times 10^{-16} \text{ Cm}) \times 3.0 \times 10^2 \text{ NC}^{-1}) \sin 30^\circ \\ &= \frac{72}{2} \times 10^{-14} \text{ Nm} \\ &= 36 \times 10^{-14} \text{ Nm}\end{aligned}$$

If a dipole is placed in a non-uniform electric field, the forces on the charges  $-q$  and  $+q$  will be unequal. Such an electric field will not only tend to rotate but also displace the dipole in the direction of the field.

### Electric Lines of Force (Field Lines)

A very convenient method for depicting the electric field (or force) is to draw lines of force pointing in the direction of the field. The sketch of the electric field lines gives us an idea of the magnitude and direction of the electric field. **The number of field lines passing through a unit area of a plane placed perpendicular to the direction of the field is proportional to the strength of the field.** A tangent at any point on the field lines gives the direction of the field at that point.

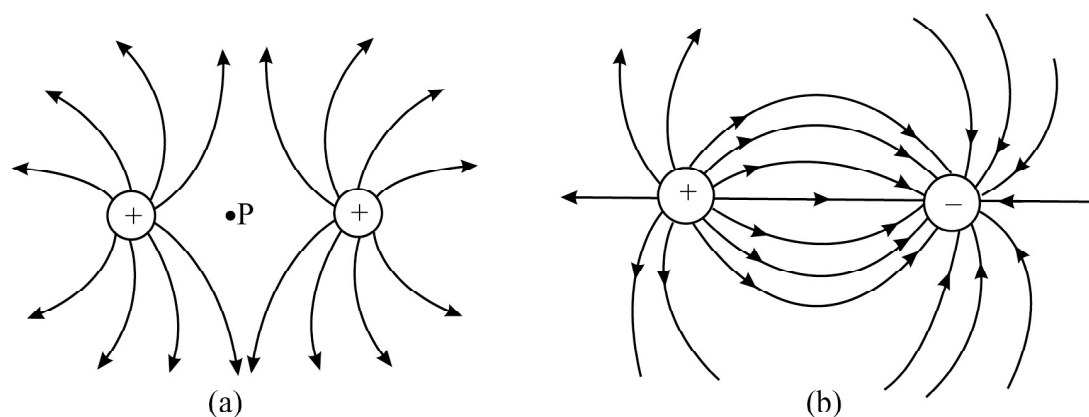
Note that the *electric field lines are only fictitious construction to depict the field. No such lines really exist.* But the behaviour of charges in the field and the interaction between charges can be effectively explained in terms of field lines. Some illustrative examples of electric field lines due to point charges are shown in Fig. 17.14. The field lines of a stationary positive charge point radially in outward direction. But for stationary negative charge, the lines start from infinity and terminate on the point charge in radially inward direction (towards the point charge). You must understand that the electric field lines in both cases are in all directions in the space. Only those which are in the plane containing the charge are shown here.



**Fig. 17.14 :** Electrical field lines of single point charges :  
 (a) The field lines of positive charge, and  
 (b) the field lines of negative charge.

Fig. 17.15 (a) shows a sketch of electric field lines of two equal and similar positive charges placed close to each other. The lines are almost radial at points very close to the positive charges and repel each other, bending outwards. There is a point P midway between the charges where no lines are present. The fields of the two charges at this point cancel each other and the resultant field at this point is zero.

Fig. 17.15 (b) depicts the field lines due to a dipole. The number of lines leaving the positive charge is equal to the number of lines terminating on the negative charge.



**Fig. 17.15 :** Electric field lines due to a system of two point charges :

- (a) Two positive charges at rest, and  
 (b) The field lines due to a dipole start from the positive charge and terminate on the negative charge.

You must remember the following properties of the electric field lines :

- The field lines start from a positive charge radially outward in all directions and terminate at infinity.
- The field lines start from infinity and terminate radially on a negative charge.
- For a dipole, field lines start from the positive charge and terminate on the negative charge.
- A tangent at any point on field line gives the direction of electric field at that point.
- The number of field lines passing through unit area of a surface drawn perpendicular to the field lines is proportional to the field strength on this surface.
- Two field lines never cross each other.

## 17.4) ELECTRIC FLUX AND GAUSS' LAW

Let us consider a sphere of radius  $r$  having charge  $+q$  located at its center. The magnitude of electric field at every point on the surface of this sphere is given by

$$E = k \times \frac{q}{r^2}$$

The direction of the electric field is normal to the surface and points outward. Let us consider a small element of area  $\Delta s$  on the spherical surface.  $\Delta \mathbf{s}$  is a vector whose magnitude is equal to the element of area  $\Delta s$  and its direction is perpendicular to this element (Fig. 17.16). The electric flux  $\Delta\phi$  is defined as the scalar product of  $\Delta \mathbf{s}$  and  $\mathbf{E}$ :

$$\Delta\phi = \mathbf{E} \cdot \Delta \mathbf{s}$$

The total flux over the entire spherical surface is obtained by summing all such contributions:

$$\phi_E = \sum_{\Delta s_i=1}^N \mathbf{E}_i \cdot \Delta \mathbf{s}_i \quad (17.21)$$

Since the angle between  $\mathbf{E}$  and  $\Delta \mathbf{s}$  is zero, the total flux through the spherical surface is given by

$$\phi_E = k \times \frac{q}{r^2} \sum \Delta s$$

The sum of all elements of area over the spherical surface is  $4\pi r^2$ . Hence the net flux through the spherical surface is

$$\begin{aligned} \phi_E &= k \times \frac{q}{r^2} \times 4\pi r^2 \\ &= 4\pi k \times q \end{aligned}$$

On substituting for  $k = 1/4\pi\epsilon_0$ , we get

$$\begin{aligned} \phi_E &= \frac{1}{4\pi\epsilon_0} \times 4\pi q \\ &= q/\epsilon_0 \end{aligned} \quad (17.22)$$

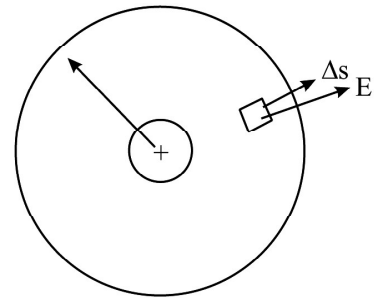


Fig. 17.16

The spherical surface of the sphere is referred to as *Gaussian surface*. Eqn. (17.22) is known as Gauss' law. It states that *the net electric flux through a closed gaussian surface is equal to the total charge  $q$  inside the surface divided by  $\epsilon_0$* .

Gauss' law is a useful tool for determining the electric field. You must also note that gaussian surface is an imaginary mathematical surface. It may not necessarily coincide with any real surface.

### Electric Field due to a Point Charge

Let us apply Gauss' law to calculate electric field due to a point charge. Draw a spherical surface of radius  $r$  with a point charge at the centre of the sphere, as shown in Fig.17.17.

The electric field  $\mathbf{E}$  is along the radial direction pointing away from the centre and normal to the surface of the sphere at every point. The normal to the element of area  $\Delta s$  is parallel to  $\mathbf{E}$ . According to Gauss' law, we can write

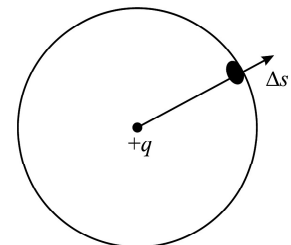


Fig. 17.17 : Electric field on a spherical surface due to a charge  $+q$  at its centre

$$\phi_E = \sum_{i=1}^N \mathbf{E}_i \cdot \Delta \mathbf{s}_i = q/\epsilon_0$$

Since  $\cos \theta = 1$  and  $\mathbf{E}$  is same on all points on the surface, we can write

$$\phi_E = E \times 4\pi r^2$$

or 
$$q/\epsilon_0 = E \times 4\pi r^2$$

$$\Rightarrow E = \frac{q}{4\pi\epsilon_0 r^2} \quad (17.23)$$

If there is a second charge  $q_0$  placed at a point on the surface of the sphere, the magnitude of force on this charge would be

$$F = q_0 \times E$$

so that

$$F = \frac{qq_0}{4\pi\epsilon_0 r^2} \quad (17.24)$$

Do you recognise this result? It is expression for Coulomb's force between two static point charges.

### Carl Friedrich Gauss (1777 – 1855)

German genius in the field of physics and mathematics, Gauss has been one of the most influential mathematicians. He contributed in such diverse fields as optics, electricity and magnetism, astronomy, number theory, differential geometry, and mathematical analysis.

As child prodigy, Gauss corrected an error in his father's accounts when he was only three year old. In primary school, he stunned his teacher by adding the integers 1 to 100 within a second.

Though he shun interactions with scientific community and disliked teaching, many of his students rose to become top class mathematicians – Richard Dedekind, Berhard Riemann, Friedrich Bessel and Sophie Germain are a few among them. Germany issued three postal stamps and a 10 mark bank note in his honour. A crater on moon called Gauss crater, and asteroid 100 called Gaussia have been named after him.



### Electric Field due to a Long Line Charge

A line charge is in the form of a **thin charged wire of infinite length** with a uniform linear charge density  $\sigma_l$  (charge per unit length). Let there be a charge  $+q$  on the wire. We have to calculate the electric field at a point P at a distance  $r$ . Draw a right circular cylinder of radius  $r$  with the long wire as the axis of the cylinder. The cylinder is closed at both ends. The surface of this cylinder is the gaussian surface and shown in Fig. 17.18. The magnitude of the electric field  $E$  is same at every point on the curved surface of the cylinder

because all points are at the same distance from the charged wire. The electric field direction and the normal to area element  $\Delta s$  are parallel.

Let the length of the gaussian cylinder be  $l$ . The total charge enclosed in the cylinder is  $q = \sigma_l$ . The area of the curved surface of the cylinder is  $2 \pi r l$ .

For the flat surfaces at the top and bottom of the cylinder, the normals to these areas are perpendicular to the electric field ( $\cos 90^\circ = 0$ ). These surfaces, therefore, do not contribute to the total flux. Hence

$$\begin{aligned} \phi_E &= \Sigma \mathbf{E} \cdot \Delta \mathbf{s} \\ &= E \times 2 \pi r l \end{aligned}$$

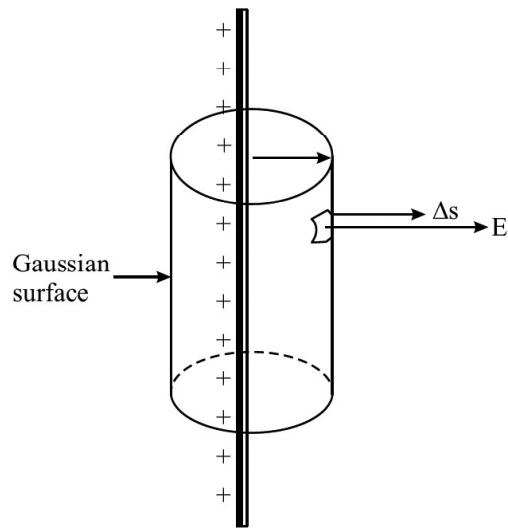
According to Gauss' law,  $\phi_E = q/\epsilon_0$ .

Hence

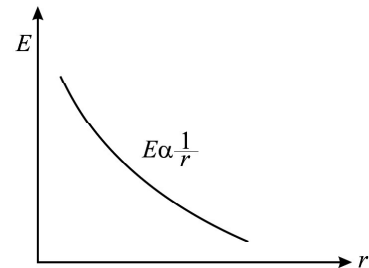
$$E \times 2 \pi r l = q/\epsilon_0 = \sigma_l/\epsilon_0$$

or 
$$E = \frac{\sigma_l}{2\pi\epsilon_0 r} \quad (17.25)$$

This shows that electric field varies inversely with distance. This is illustrated in Fig.17.19.



**Fig. 17.18 :** Electric field due to an infinite line of charges having uniform linear charge density. The gaussian surface is a right circular cylinder.



**Fig. 17.19 :** Variation of  $E$  with  $r$  for a line charge

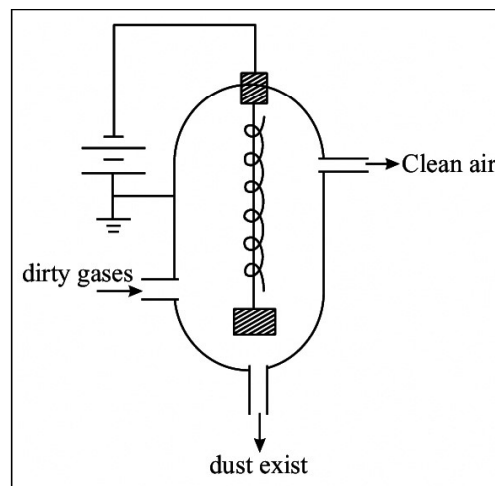
### Intext Questions 17.4

- If the electric flux through a gaussian surface is zero, does it necessarily mean that
  - the total charge inside the surface is zero?
  - the electric field is zero at every point on the surface?
  - the electric field lines entering into the surface is equal to the number going out of the surface?
- If the electric field exceeds the value  $3.0 \times 10^6 \text{ NC}^{-1}$ , there will be sparking in air. What is the highest value of charge that a metallic sphere can hold without sparking in the surrounding air, if the radius of the sphere is 5.0 cm?
- What is the magnitude and direction of the net force and net torque on a dipole placed along a
  - uniform electric field, and
  - non-uniform field.

### Electrostatic Filter

You must have seen black smoke and dirt particles coming out of a chimney of a thermal power station or brick kiln. The smoke consists of not only gases but large quantities of small dust (coal) particles. The smoke along with the dirt is discharged into the atmosphere. The dust particles settle down on earth and pollute the soil. The gases contribute to global warming. These are extremely injurious to living systems (health). It is therefore essential that the dirt is removed from smoke before it is discharged into the atmosphere.

A very important application of electrical discharge in gases by application of high electric field is the construction of a device called *Electrostatic Filter or Precipitator*.



The basic diagram of the device is shown here. The central wire inside a metallic container is maintained at a very high negative potential (about 100 kV). The wall of the container is connected to the positive terminal of a high volt battery and is **earthed**. A weight  $W$  keeps the wire straight in the central part. The electric field thus created is from the wall towards the wire. The dirt and gases are passed through the container. An electrical discharge takes place because of the high field near the wire. Positive and negative ions and electrons are generated. These negatively charged particles are accelerated towards the wall. They collide with dust particles and charge them. Most of the dust particles become negatively charged because they capture electrons or negative ions. They are attracted towards the wall of the container. The container is periodically shaken so that the particles leave the surface and fall down at the bottom of the container. These are taken out through the exit pipe.

The undesirable dust particles are thus removed from the gases and the clean air goes out in the atmosphere. Most efficient systems of this kind are able to remove about 98% of the ash and dust from the smoke.

### WHAT YOU HAVE LEARNT

- Electric charge is produced when glass rod is rubbed with silk or rubber is rubbed with fur.
- By convention, the charge on glass rod is taken **positive** and that on rubber is taken negative.
- Like charges repel and unlike charges attract each other.



- Coulomb's law gives the magnitude and direction of force between two point charges:

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

$$\text{where } k = \frac{1}{4\pi\epsilon_0} = 9.0 \times 10^9 \text{ Nm}^2\text{C}^{-2}$$

- The smallest unit of charge in nature is the charge on an electron :

$$e = 1.60 \times 10^{-19} \text{ C}$$

- Charge is conserved and quantised in terms of electronic charge.
- The electric field  $\mathbf{E}$  due to a charge  $q$  at a point in space is defined as the force experienced by a unit test charge  $q_0$  :

$$\mathbf{E} = \mathbf{F}/q_0 = k \times \frac{q}{r^2} \hat{\mathbf{r}}$$

- Superposition principle can be used to obtain the force experienced by a charge due to a group of charges. It is also applicable to electric field at a point due to a group of charges.
- Electric dipole is a system of two equal and unlike charges separated by a small distance. It has a dipole moment  $|\mathbf{p}| = qr$ ; the direction of  $\mathbf{p}$  is from negative charge to positive charge along the line joining the two charges.
- The electric field due to a dipole in end-on position and broad-on position is respectively given by

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{2\mathbf{p}}{r^3}$$

and

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{\mathbf{p}}{r^3}.$$

- Electric field lines (line of force) are only a pictorial way of depicting field.
- Electric flux is the total number of electric lines of force passing through an area and is defined as  $\phi_E = \mathbf{E} \cdot \mathbf{A}$ .
- Gauss's law states that the total flux passing through a closed area is  $\frac{1}{\epsilon_0}$  times the total charge enclosed by it.
- The electric field due to a line charge is given by  $E = \frac{\sigma_l}{2\pi\epsilon_0 r}$ .

<b>TERMINAL EXERCISE</b>
--------------------------

1. A  $+12\mu\text{C}$  charge is at  $x = 20\text{ cm}$  and a  $-18\mu\text{C}$  ( $-q$ ) charge is at  $x = 29\text{ cm}$  on the  $x$ -axis. Calculate the magnitude and direction of the force on a charge of  $18\mu\text{C}$ . What is the direction of force on  $12\mu\text{C}$  charge?
2. Two point Charges  $q_1$  and  $q_2$  separated by a distance of  $3.0\text{ m}$  experience a mutual force of  $16 \times 10^{-15}\text{ N}$ . Calculate the magnitude of force when the separation distance is changed to  $6.0\text{ m}$ .
3. There are two points A and B separated by a distance  $x$ . If two point charges  $+q$  each are on the points A and B, the force between them is  $F$ . The point charges are now replaced by two identical metallic spheres having the same charge  $+q$  on each. The distance between their centers is again  $x$  only. Will the force between them change? Give reasons to support your answer.
4. Two alpha particles are separated by a distance of  $10^{-13}\text{ m}$ . Calculate the force between them in free space.
5. The force of repulsion between two point chargers placed  $16\text{ cm}$  apart in vaccum is  $7.5 \times 10^{-10}\text{ N}$ . What will be force between them, if they are placed in a medium of dielectric constant  $k = 2.5$ ?
6. Compare the electrical force with the gravitational force between two protons separated by a distance  $x$ . Take charge on proton as  $1.60 \times 10^{-19}\text{ C}$ , mass of proton as  $1.67 \times 10^{-27}\text{ kg}$  and Gravitational constant  $G = 6.67 \times 10^{-11}\text{ Nm}^2\text{kg}^{-2}$ .
7. Four identical point charges  $+q$  each are placed at the four corners (one  $q$  at one corner) of a square of side  $l$ . Find the force experienced by a test charge  $q_0$  placed at the center of the square.
8. When are the electric field lines parallel to each other?
9. How many electrons should be removed from a metallic sphere to give it a positive charge  $= 6.4 \times 10^{-7}\text{C}$ .
10. Consider an electric dipole of  $q = 3.0 \times 10^{-6}\text{ C}$  and  $2l = 4 \times 10^{-10}\text{ m}$ . Calculate the magnitude of dipole moment. Calculate electric field at a point  $r = 6 \times 10^{-6}\text{m}$  on the equatorial plane.
11. A Charge  $-q = 15 \times 10^{-6}\text{ C}$  is placed on a metallic sphere of radius  $R = 3.0\text{ mm}$ . Calculate the magnitude and direction of the electric field at a point  $r = 15\text{ cm}$  from the center of the sphere. What will be the magnitude and direction of the field at the same point if  $3.0\text{ mm}$  sphere is replaced by  $9.0\text{ mm}$  sphere having the same Charge.
12. A charge of  $+15\mu\text{C}$  is located at the center of a sphere of radius  $20\text{ cm}$ . Calculate the electric flux through the surface of the sphere.
13. A proton is placed in a uniform electric field  $E = 8.0 \times 10^4\text{ NC}^{-1}$ . Calculate the acceleration of the proton.

14. Two point charges  $q_1$  and  $q_2$  are 3.0 cm apart and  $(q_1 + q_2) = 20 \mu\text{C}$ . If the force of repulsion between them is 750 N, calculate  $q_1$  and  $q_2$ .

### ANSWERS TO INTEXT QUESTIONS

#### 17.1

- (i) Yes (ii) Charge =  $3.2 \times 10^{-17} \text{ C}$  (Negative).
- A has charge + Q. When A and B are brought in contact, charge will get distributed equally.  
(i) Yes., (ii) + Q/2
- $q = 4.8 \times 10^{-16}$

Since  $Ne = q$ , we get

$$N = \frac{4.8 \times 10^{-16}}{1.6 \times 10^{-19}} = 3.0 \times 10^3 \text{ charges}$$

#### 17.2

- $Q_1 = 16 \mu\text{C}$ ,  $Q_2 = 9 \mu\text{C}$  and  $r = 12\text{m}$

Since

$$F = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r^2}$$

$$= \frac{(9 \times 10^9 \text{ Nm}^2 \text{C}^{-2})(16 \times 10^{-6} \text{ C})(9 \times 10^{-6} \text{ C})}{144 \text{ m}^2} = 9 \times 10^{-3} \text{ N}$$

(i) direction from  $q_2$  to  $q_1$

(ii) direction from  $q_1$  to  $q_2$

- The force at A due to charge at B,  $F_1 = k \frac{q^2}{a^2}$  where  $AB = a$

Since  $AB = AC$ , the force at A due to charge at B is

$$F_2 = k \frac{q^2}{a^2}$$

$$R^2 = F_1^2 + F_2^2 = 2F^2$$

$$R = F\sqrt{2} \text{ at } 45^\circ$$

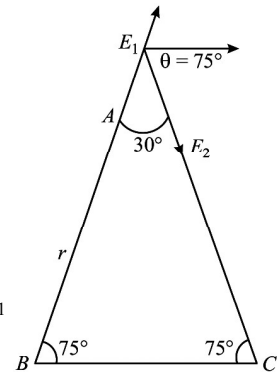
## 17.3

- (a)  $E$  along the + x axis.
  - (b) along the + y axis.
  - (c) at  $45^\circ$  with the x axis
- $AB = AC = 40 \text{ cm}$

$$|\mathbf{E}_1| = \frac{kq}{r^2} = |\mathbf{E}_2| = \frac{9 \times 10^9 \text{ Nm}^2 \text{ C}^{-2} \times (2 \times 10^{-6} \text{ C})}{(0.40 \text{ m})^2} = 1.125 \times 10^5 \text{ NC}^{-1}$$

The resultant of  $\mathbf{E}_1$  and  $\mathbf{E}_2$  will be parallel to BC. Hence

$$\begin{aligned} R^2 &= E_1^2 + E_2^2 + 2E_1 E_2 \cos 150 \\ &= 2 E^2 + 2 E^2 \cos (180 - 30) \\ &= 2 E^2 - 2 E^2 \times \cos 30 = 2 E^2 \left( 1 - \frac{\sqrt{3}}{2} \right) = 4.723 \times 10^{10} \text{ N}^2 \text{ C}^{-2} \end{aligned}$$



Direction will be parallel to BC in the direction  $B \rightarrow C$ .

- $E$  is directed towards the earth. The force on  $-ve$  charge will be vertically upwards.
- The field will be zero at the mid point between the charge.

## 17.4

- (a) Yes (b) not necessarily (c) Yes.

$$2. \quad E = \frac{Q}{4\pi\epsilon_0 r^2}$$

$$\therefore Q = 4\pi\epsilon_0 r^2 E$$

$$= (3 \times 10^6 \text{ NC}^{-1}) \times \frac{1}{(9 \times 10^9 \text{ Nm}^2 \text{ C}^{-2})} \times (25 \times 10^{-4} \text{ m}^2)$$

$$= 8.3 \times 10^{-7} \text{ C}$$

- (a)  $\mathbf{F} = 0, \tau = 0$
  - (b)  $\mathbf{F} \neq 0, \tau = 0$

**ANSWERS TO TERMINAL EXERCISE**

1. 240N towards negative x-direction. Force on +12  $\mu\text{C}$  charge is towards + x direction.
2.  $F = 4 \times 10^{-15} \text{ N}$
4.  $92.39 \times 10^{-3} \text{ N}$
5.  $3 \times 10^{-10} \text{ N}$
6. Electric force is  $10^{36}$  times the gravitational force.
7. zero.
9.  $4 \times 10^{12}$  electrons
10.  $12 \times 10^{-16} \text{ Cm}$ .  $0.5 \times 10^{11} \text{ Nc}^{-1}$
11.  $6 \times 10^6 \text{ NC}^{-1}$  towards the centre, same field.
12.  $1.7 \times 10^6 \text{ C}^{-1} \text{ N m}^2$
13.  $7.6 \times 10^{12} \text{ ms}^{-2}$
14.  $15\mu\text{C}$  and  $5 \mu\text{C}$ .



# 18

## ELECTROSTATIC POTENTIAL AND CAPACITANCE

### INTRODUCTION

In previous chapters, you learnt about the direction of flow of fluids and thermal energy. You may recall that the flow of water is determined by the relative heights of water in two interconnected containers. If the level of water in one container is higher than that in the other, water will flow from higher level to lower level, irrespective of the quantity of water in the containers. Temperature plays a similar role in case of flow of thermal energy from one object to another. Thermal energy always flows from a body at higher temperature to the one at lower temperature. Here also, the direction of flow does not depend on the quantity of thermal energy possessed by an object.

Electric potential plays a similar role in the flow of charges from one point to another. The positive charge always moves from a point at higher potential to a point at lower potential. A positive test charge, when left free in an electric field, it moves in the direction of the electric field. From this behaviour of a positive test charge, you may be tempted to say that the **electric field (E)** and **electric potential (V)** are closely related. In this lesson, you will learn to establish a relation between these physical quantities. You will also learn about a device called capacitor, which is used to store the charge, filter alternating current and finds wide applications in electronic circuitry as well in power transmission.

### OBJECTIVES

After studying this lesson, you should be able to

- *explain the meaning of electric potential at a point and potential difference;*
- *derive an expressions for electric potential due to a point charge and a dipole;*
- *explain the principle of capacitors and their applications;*
- *derive an expression for the capacitance of a parallel plate capacitor;*
- *obtain equivalent capacitance in grouping of capacitors;*
- *calculate the energy stored in a capacitor;*
- *explain polarization of dielectric materials in an electric field.*

### 18.1 ELECTRIC POTENTIAL AND POTENTIAL DIFFERENCE

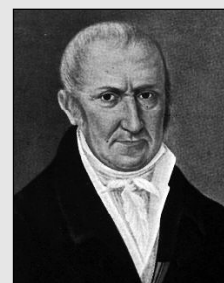
When a charged particle is made to move in an electrostatic field in a direction opposite to the direction of the field, work is done by an external agency. This work is stored as

potential energy of charge in accordance with the law of conservation of energy. So we can say that an electric charge placed at a point in an electric field has potential energy, which is a function of its position. We can visualize the potential energy of charge in the field as a scalar function of position and for a unit charge it is called as potential. It means that different points in an electric field would be at different potentials. And if a positively charged particle is placed in an electric field, it will tend to move from higher to lower potential to minimize its potential energy. In the next lesson, you will learn how the concept of potential difference leads to the flow of current in electric circuits.

*The **electric potential** at any point in an electric field is equal to the work done against the electric force in moving a unit positive charge from infinity to that point.* Electric potential is a scalar quantity, as it is related to work done.

### Alessandro, Conte Volta (1745-1827)

Born at Como, Italy, Volta was a professor at Pavia for more than 20 years. A well travelled man, he was known to many famous men of his times. He decisively proved that animal electricity observed by Luigi Galvani in frog muscles was a general phenomenon taking place between two dissimilar metals separated by acidic or salt solutions. On the basis of this observation, he invented first electro-chemical cell, called voltaic cell. The unit of potential difference is named volt in his honour.



The potential at a point is taken positive when work is done against the field by a positive charge but negative when work is done by the electric field in moving the unit positive charge from infinity to the point in the field.

Consider two points A and B in an electric field (Fig. 18.1). If a test charge  $q_0$  is moved from point A to point B along any path by an external force, the amount of work done by the external force is given by

$$W_{AB} = q_0 (V_B - V_A) \quad (18.1)$$

Thus, potential difference between points A and B will be

$$V_{AB} = V_B - V_A = \frac{W_{AB}}{q_0} \quad (18.2)$$

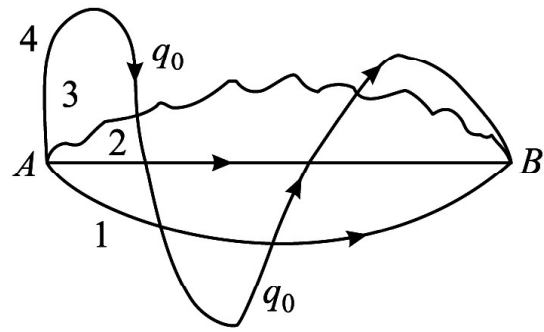
Here  $V_A$  and  $V_B$  are potentials at points A and B, respectively.

A potential difference is said to exist between two points in an electric field, if work is done against the electric force in moving a positive test charge from one point to the other. Note that this work is independent of the path. (*For this reason, the electric field is said to be a conservative field*). The SI unit of potential and potential difference is **volt** :

$$1 \text{ volt} = 1 \text{ joule/1 coulomb}$$

If one joule of work is done in taking a test charge of one coulomb from one point to the other in an electric field, the potential difference between these points is said to be one volt. If one joule of work is done in bringing a test charge of one coulomb from infinity to a point in the field, the potential at that point is one volt.

Note that potential at a point is not a unique quantity as its value depends on our choice of zero potential energy (infinity). However, the potential difference between two points in a stationary field will have a unique value. Let us now learn to calculate potential at a point due to a single charge.

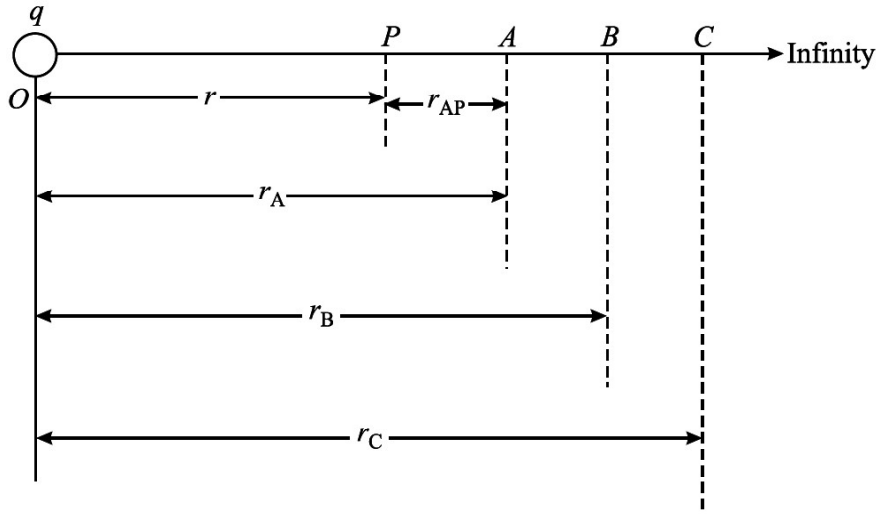


**Fig. 18.1 :** The work done in moving a test charge from one point to another in an electric field is independent of the path followed.

### Potential at a point due to a Point Charge

Suppose we have to calculate electric potential at point P due to a single point charge  $+q$  situated at O (Fig. 18.2), where  $OP = r$ . The magnitude of electric field at P due to the point charge is given by

$$E_P = \frac{1}{4\pi\epsilon_0} \times \frac{q}{r^2} \quad (18.3)$$



**Fig. 18.2 :** Work done per unit charge in moving a charge  $q_0$  from infinity to a point P in an electric field E is the potential at that point.

Similarly, the electric field at point A will be

$$E_A = \frac{1}{4\pi\epsilon_0} \frac{q}{r_A^2} \quad (18.4)$$



If points P and A are very close, the average field  $E_{AP}$  between these points can be taken as the geometric mean of  $E_P$  and  $E_A$  :

$$\begin{aligned} E_{AP} &= \sqrt{E_A \times E_P} \\ &= \sqrt{\frac{1}{4\pi\epsilon_0} \frac{q}{r_A^2} \times \frac{1}{4\pi\epsilon_0} \frac{q}{r^2}} \\ &= \frac{1}{4\pi\epsilon_0} \frac{q}{r_A r} \end{aligned} \quad (18.5)$$

Therefore, the magnitude of force experienced by a test charge  $q_0$  over this region will be

$$F_{AP} = q_0 E_{AP} = \frac{1}{4\pi\epsilon_0} \frac{q q_0}{r_A r} \quad (18.6)$$

and the work done in moving charge  $q_0$  from A to P is given by

$$\begin{aligned} W_{AP} &= F_{AP} \times r_{AP} \\ &= \frac{1}{4\pi\epsilon_0} \frac{q q_0}{r_A r} \times (r_A - r) \\ &= \frac{q q_0}{4\pi\epsilon_0} \times \left( \frac{1}{r} - \frac{1}{r_A} \right) \end{aligned} \quad (18.7)$$

where  $r_{AP}$  is the distance between points A and P.

Similarly, work done in moving this charge from B to A will be given by

$$W_{BA} = \frac{q q_0}{4\pi\epsilon_0} \times \left( \frac{1}{r_A} - \frac{1}{r_B} \right) \quad (18.8 \text{ a})$$

And work done in moving the test charge from C to B will be

$$W_{CB} = \frac{q q_0}{4\pi\epsilon_0} \times \left( \frac{1}{r_B} - \frac{1}{r_C} \right) \quad (18.8 \text{ b})$$

and so on. The total work done in moving the charge from infinity to the point P will be

$$W = \frac{q q_0}{4\pi\epsilon_0} \times \left( \frac{1}{r} - \frac{1}{r_A} + \frac{1}{r_A} - \frac{1}{r_B} + \frac{1}{r_B} - \frac{1}{r_C} \dots + \dots - \frac{1}{\infty} \right)$$

$$\begin{aligned}
 &= \frac{q q_0}{4\pi\epsilon_0} \times \left( \frac{1}{r} - \frac{1}{\infty} \right) \\
 &= \frac{q q_0}{4\pi\epsilon_0 r} \qquad (18.9)
 \end{aligned}$$

By definition, potential at a point is given by

$$\begin{aligned}
 V_p &= \frac{W}{q_0} \\
 &= \frac{1}{4\pi\epsilon_0} \frac{q}{r} \qquad (18.10)
 \end{aligned}$$

Note that potential is inversely proportional to distance. It is positive or negative depending on whether  $q$  is positive or negative.

If there are several charges of magnitudes  $q_1, q_2, q_3, \dots$ , the electric potential at a point is the scalar sum of the potentials due to individual charges (Fig. 18.3) :

$$\begin{aligned}
 V &= V_1 + V_2 + V_3 + \dots \\
 &= \sum_{i=1}^{\infty} \frac{q_i}{4\pi\epsilon_0 r_i} \qquad (18.11)
 \end{aligned}$$

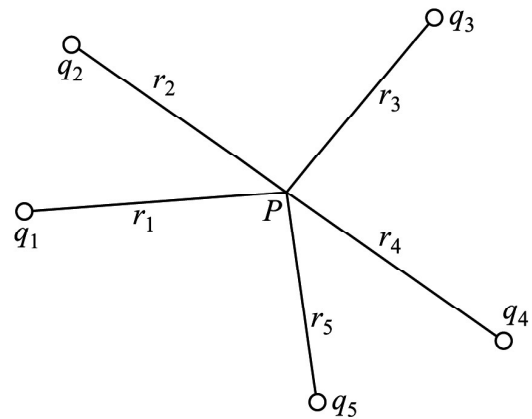


Fig. 18.3 : Potential at a point P due to a system of charges

### Potential at a Point due to an Electric Dipole

Let us consider an electric dipole consisting of two equal and opposite point charges  $-q$  at A and  $+q$  at B, separated by a distance  $2l$  with centre at O. We wish to calculate potential at a point P, whose polar co-ordinates are  $(r, \theta)$ ; i.e.  $OP = r$  and  $\angle BOP = \theta$ , as shown in Fig.18.4. Here  $AP = r_1$  and  $BP = r_2$ . We can easily calculate potential at P due to point charges at A and B using Eqn.(18.10) :

$$V_1 = \frac{1}{4\pi\epsilon_0} \times \frac{(-q)}{r_1}$$

and

$$V_2 = \frac{1}{4\pi\epsilon_0} \times \frac{q}{r_2}$$

Total potential at P due to both the charges of the dipole is given by

$$V = V_1 + V_2$$

That is,

$$V = \frac{q}{4\pi\epsilon_0} \left[ \frac{1}{r_2} - \frac{1}{r_1} \right] \quad (18.12)$$

To put this result in a more convenient form, we draw normals from A and B on the line joining O and P. From  $\Delta BOD$ , we note that  $OD = l \cos \theta$  and from  $\Delta OAC$  we can write  $OC = l \cos \theta$ . For a small dipole ( $AB \ll OP$ ), from Fig. 18.4, we can take  $PB = PD$  and

$PA = PC$ . Hence

$$r_1 = r + l \cos \theta$$

$$r_2 = r - l \cos \theta$$

Using these results in Eqn (18.12), we get

$$\begin{aligned} V &= \frac{q}{4\pi\epsilon_0} \left[ \frac{1}{(r-l\cos\theta)} - \frac{1}{(r+l\cos\theta)} \right] \\ &= \frac{q}{4\pi\epsilon_0} \left[ \frac{(2l\cos\theta)}{(r^2-l^2\cos^2\theta)} \right] \\ &= \frac{q \times 2l \cos \theta}{4\pi\epsilon_0 r^2} \end{aligned}$$

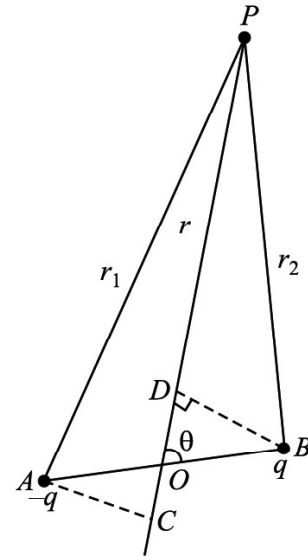


Fig. 18.4 : Electric potential at a point P due to an electric dipole.

Here we have neglected the term containing second power of  $l$  since  $l \ll r$ .

In terms of dipole moment ( $\mathbf{p} = q \times 2l$ ), we can express this result as

$$V = \frac{\mathbf{p} \cos \theta}{4\pi\epsilon_0 r^2} \quad (18.13)$$

This result shows that unlike the potential due to a point charge, the potential due to a dipole is inversely proportional to the square of the distance.

### Special Cases

**Case I :** When point P lies on the axial line of the dipole on the side of positive charge  $q$  ( $\theta = 0$ ) and  $\cos 0 = 1$ . Then Eqn. (18.13) reduces to

$$V_{\text{AXIS}} = \frac{\mathbf{p}}{4\pi\epsilon_0 r^2} \quad (18.14)$$

**Case II :** When point P lies on the axial line of the dipole but on the side of negative charge  $q$  ( $\theta = 180^\circ$ ) and  $\cos 180^\circ = -1$ . Hence

$$V_{\text{AXIS}} = -\frac{\mathbf{p}}{4\pi\epsilon_0 r^2} \quad (18.15)$$

**Case III :** When point P lies on the equatorial line of the dipole (perpendicular bisector of AB),  $\theta = 90^\circ$  and  $\cos 90^\circ = 0$ . Then

$$V_{\text{equatorial}} = 0 \quad (18.16)$$

That is, electric potential due to a dipole is zero at every point on the equatorial line of the dipole. The potential at all points in this plane will be same, i.e. zero. Such a surface is referred to as *equipotential surface*. The electric field is always perpendicular to an equipotential surface. No work is done in moving a charge from one point to another on the equipotential surface.

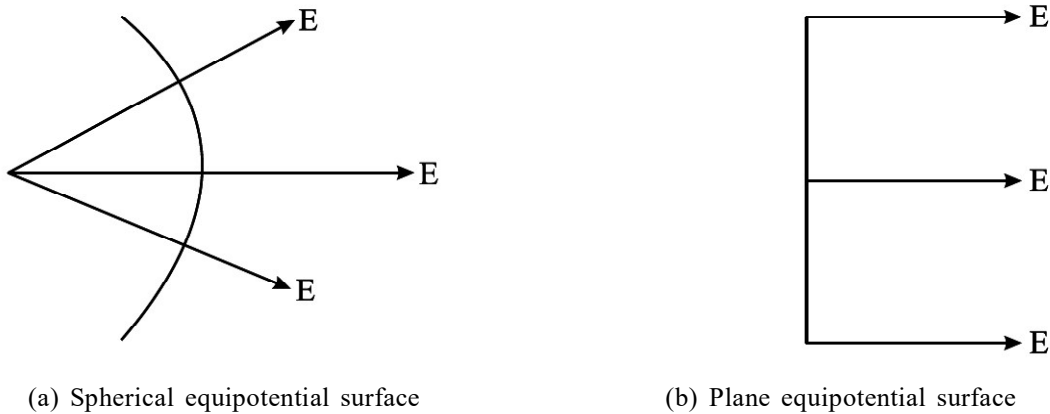


Fig. 18.5 : Equipotential surfaces and electric field directions

### Potential Energy of a System of Point Charges

The electric potential energy is the energy possessed by a system of point charges by virtue of their being in an electric field. When charges are infinite distance apart, they do not interact and their potential energy is zero. If we want to assemble a charge system, i.e. bring charges near each other, work will have to be done. This work is stored in the form of potential energy in the system of these charges. This is called the electric potential energy of the charge system. Hence, we can define *potential energy of a system of point charges as the total amount of work done in bringing various point charges of the system to their respective positions from infinitely large mutual separations.*

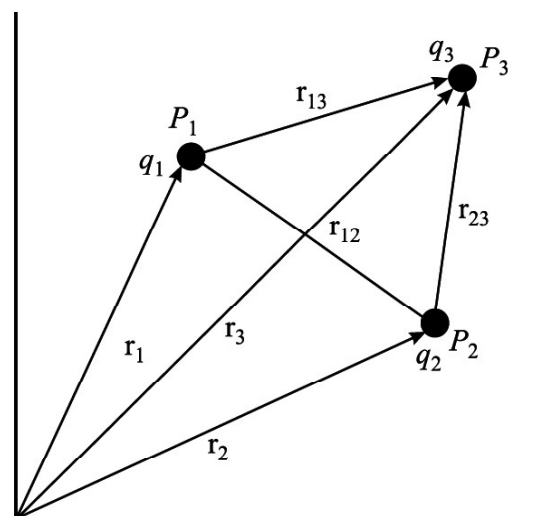


Fig. 18.6 : Potential energy of a system of point charges separated by a distance

Suppose that a point charge  $q_1$  is located at a point  $P_1$  with position vector  $\mathbf{r}_1$  in space. Assume that point charge  $q_2$  is at infinity.

This has to be brought to the point  $P_2$  having position vector  $\mathbf{r}_2$  where  $P_1P_2 = \mathbf{r}_{12}$ , as shown in Fig.18.6. We know that electric potential at  $P_2$  due to charge  $q_1$  at  $P_1$  is

$$V_2 = \frac{1}{4\pi\epsilon_0} \frac{q_1}{|\mathbf{r}_{12}|} \quad (18.17)$$

From the definition of potential, work done in bringing charge  $q_2$  from infinity to point  $P_2$  is

$$W = (\text{Potential at } P_2) \times \text{value of charge}$$

This work is stored in the system of charges  $q_1$  and  $q_2$  in the form of electric potential energy  $U$ . Thus,

$$U = \frac{q_1 \times q_2}{4\pi\epsilon_0 |\mathbf{r}_{12}|} \quad (18.18)$$

In this case the two charges have the same sign, hence work is done against the repulsive force to bring them closer and hence, electric potential energy of the system increases. On the other hand, in separating them from one another, work is done by the field. As a result the potential energy of the system decreases. If charges are of opposite sign, i.e. one is positive and the other is negative, the potential energy of the charge system decreases in bringing the charges closer and increases in separating them from one another.

For a three point charge system (Fig. 18.6), Eqn. (18.18) can be written as

$$U = \frac{1}{4\pi\epsilon_0} \left( \frac{q_1q_2}{r_{12}} + \frac{q_1q_3}{r_{13}} + \frac{q_2q_3}{r_{23}} \right) \quad (18.19)$$

Proceeding in the same way, we can calculate the potential energy of a system of any number of charges.

By combining Eqns. (18.3) and (18.13), the potential energy of a dipole in a uniform electric field can be written as

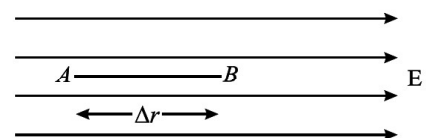
$$Uq = - \mathbf{p}E \cos\theta = - \mathbf{p} \cdot \mathbf{E} \quad (18.20)$$

where  $\mathbf{p}$  is the dipole moment in electric field  $\mathbf{E}$  and  $\theta$  is the angle between  $\mathbf{p}$  and  $\mathbf{E}$ .

## 18.2) RELATION BETWEEN ELECTRIC FIELD AND POTENTIAL

Consider two points A and B in a uniform electric field  $\mathbf{E}$ , separated by a small distance  $\Delta r$ . By definition, potential difference  $\Delta V$  between A and B is equal to the work done in moving a unit positive test charge from A to B:

$$\begin{aligned} \Delta V &= (\text{Force on unit positive charge}) \times (AB) \\ &= \mathbf{E} \cdot \Delta \mathbf{r} = E (\Delta r) \cos 180^\circ \end{aligned}$$



$$= - \mathbf{E} \Delta r$$

$$\text{or} \quad \mathbf{E} = - \frac{\Delta V}{\Delta r} \quad (18.21)$$

The negative sign indicates that work is done against the electric field.

Hence, at any point, the electric field is equal to the ratio of negative change of potential with distance (called ***potential gradient***) at that point in the direction of field. Remember that electric potential is a scalar quantity but electric potential gradient is a vector as it is numerically equal to electric field.

From the above relation, for a uniform electric field, we can write

$$E = \frac{V_A - V_B}{d} \quad (18.22)$$

Here  $V_A$  and  $V_B$  are potentials at points A and B, respectively separated by a distance  $d$ .

### Example 18.1

In a 10 volt battery, how much work is done when a positively charged particle having charge  $1.6 \times 10^{-19}$  C is moved from its negative terminal to the positive terminal?

**Solution :**

According to Eqn. (18.2)

$$V_{AB} = W_{AB} / q_0$$

Since  $V_{AB} = 10$  V and  $q_0 = 1.6 \times 10^{-19}$  C, we get

$$\begin{aligned} W_{AB} &= (10\text{V}) \times (1.6 \times 10^{-19}\text{C}) \\ &= 1.6 \times 10^{-18} \text{ J} \end{aligned}$$

### Example 18.2

A point charge  $q$  is at the origin of Cartesian co-ordinate system. The electric potential is 400 V and the magnitude of electric field is  $150 \text{ NC}^{-1}$  at a point  $x$ . Calculate  $x$  and  $q$ .

**Solution :**

The electric field

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

On inserting the numerical values, we get

$$150 = \frac{400}{x}$$

$$\text{or} \quad x = 2.67 \text{ m}$$

Recall that electric field is given by the expression

$$E = \frac{1}{4\pi\epsilon_0} \frac{q}{x^2}$$

We substitute  $\frac{1}{4\pi\epsilon_0} = 9 \times 10^9 \text{ NC}^{-2} \text{ m}^2$ ,  $E = 150 \text{ NC}^{-1}$  and  $x = 2.67$  and obtain

$$\begin{aligned} q &= \frac{(150 \text{ NC}^{-1}) \times (2.67 \text{ m})^2}{9 \times 10^9 \text{ NC}^{-2}} \\ &= 11.9 \times 10^{-8} \text{ C} \end{aligned}$$

### Intext Questions 18.1

1. A metallic sphere of radius  $R$  has a charge  $+q$  uniformly distributed on its surface. What is the potential at a point  $r$  ( $> R$ ) from the centre of the sphere?
2. Calculate the work done when a point charge is moved in a circle of radius  $r$  around a point charge  $q$ .
3. The electric potential  $V$  is constant in a region. What can you say about the electric field  $E$  in this region?
4. If electric field is zero at a point, will the electric potential be necessarily zero at that point.
5. Can two equipotential surfaces intersect?

On the basis of charge conduction, substances are broadly classified as **conductors** and **insulators**. In solids, conduction of electricity usually takes place due to free electrons. Conductors have free charge carriers through which electric currents can be established on applying an electric field. Metals are good conductors. Substances having no free charge carriers are called **insulators**. The common insulators are wood, ebonite, glass, quartz, mica etc. Substances which have electrical conductivity in between those of conductors and insulators are called **semiconductors**. The ratio of electrical conductivities of good conductors to good insulators is of the order of  $10^{20}$ . Let us now learn how conductors behave in an electric field.

### Behaviour of Conductors in an Electric Field

Conductors have electrons which are not bound tightly in their atoms. These are free to move within the conductor. However, there is no net transfer of electrons (charges) from one part of the conductor to the other in the absence of any applied electric field. The conductor is said to be in electrostatic equilibrium.

Refer to Fig. 18.7 (a) which shows a conductor placed in an external electric field  $\mathbf{E}$ . The free electrons are accelerated in a direction opposite to that of the electric field. This results in build up of electrons on the surface ABCD of the conductor. The surface FGHK becomes positively charged because of removal of electrons. These charges (-ve on surface

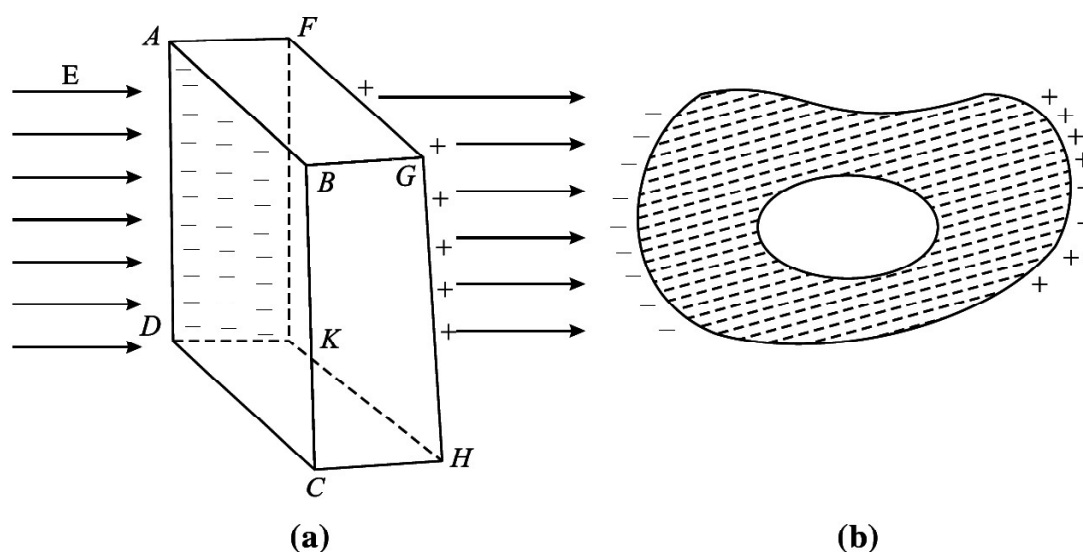
ABCD and +ve on surface FGHK) create an electric field  $E_1$  which is in a direction opposite to the applied field  $E$ . The transfer of electrons from ABCD to FGHK continues till  $E$  becomes equal to  $E_1$ . Such a state of electrostatic equilibrium is reached usually in  $10^{-16}$  s. We then say that equilibrium is reached almost instantaneously. If there is a cavity inside a conductor, the electric field inside the cavity is zero [Fig. 18.7 (b)].

These results are true for a charged conductor or when charges are induced on a neutral conductor by an external electric field.

This property of a conductor is used in **Electrostatic Shielding** — a phenomenon of protecting a certain region of space from external electric fields. To protect delicate instruments from external electric fields, they are enclosed in hollow conductors. That is why in a thunder storm accompanied by lightning, it is safer to be inside a car or a bus than outside. The metallic body of the car or bus provides electrostatic shielding from lightning.

Conductors in electrostatic equilibrium exhibit the following properties :

- There is no electric field inside a conductor.
- The electric field outside a charged conductor is perpendicular to the surface of the conductor, irrespective of the shape of the conductor.
- Any charge on the conductor resides on the surface of the conductor.



**Fig. 18.7 :** Electrostatic shielding:(a) External electric field  $E$  pulls free electrons on the surface ABCD. The surface FGHK, which is deficient in electrons, becomes positively charged; the net field inside the conductor is zero. (b) If there is a cavity inside a conductor, the field inside the cavity is zero.

### 18.3 CAPACITANCE

Let us consider two conductors having equal but opposite charges  $+Q$  and  $-Q$  on them. There is a potential difference  $V$  between them. Such a system of conductors is called a **capacitor**. Experimentally it is found that the potential difference is directly proportional to charge on a conductor. As charge increases, the potential difference between them also increases but their ratio remains constant. This ratio is termed as capacitance of the capacitor:



$$\begin{aligned}
 Q &\propto V \\
 Q &= CV \\
 C &= Q/V
 \end{aligned}
 \tag{18.23}$$

Here  $C$  is introduced as proportionality constant. Capacitance is independent of charge and potential. Capacitance depends on the area of the plates, distance between the plates of capacitor and the dielectric placed between the plates. The capacitance is defined as the ratio between the charge on either of the conductors and the potential difference between them. It is a measure of the capability of a capacitor to store charge.

In SI system of units, capacitance is measured in farad (F). The capacitance is one farad, if a charge of one coulomb creates a potential difference of one volt :

$$1 \text{ farad} = \frac{1 \text{ coulomb}}{1 \text{ volt}} \tag{18.24}$$

You may recall from the previous unit that coulomb is a very large unit of charge. It means that farad is also a very large unit of capacitance. Usually we use capacitors of values in microfarad or picofarad:

$$1 \text{ microfarad} = 10^{-6} \text{ farad, written as } \mu\text{F}$$

$$1 \text{ picofarad} = 10^{-12} \text{ farad, written as pF}$$

In an electrical circuit, a capacitor is represented by two parallel lines.

### Capacitance of a Spherical Conductor

Suppose that a sphere of radius  $r$  is given charge  $q$ . Let the potential of the sphere be  $V$ . Then

$$V = \frac{1}{4\pi\epsilon_0} \frac{q}{r}$$

Since  $C = q/V$ , we find that

$$C = \frac{q}{q/4\pi\epsilon_0 r} = 4\pi\epsilon_0 r = \frac{r}{9 \times 10^9} \tag{18.25}$$

This shows that capacitance of a spherical conductor is directly proportional to its radius. In fact, it is numerically equal to its radius divided by  $9 \times 10^9$ , where radius is taken in metre. For example, the capacitance of a sphere of radius 0.18 m is

$$C = \frac{0.18}{9} \times 10^{-9} \text{ F} = 20 \text{ pF}$$

### Working Principle of Capacitor

You will come across many types of capacitors in your physics laboratory. The power supply system of your city also uses capacitors. These also form important components of devices such as radio, T.V., amplifiers and oscillators. A capacitor essentially consists of

two conductors, one charged and the other usually earthed. To understand the principle of a capacitor, let us consider an insulated metal plate *A* and give it positive charge (*q*) till its potential (*V*) becomes maximum. (Any further charge given to it would leak out.) The capacitance of this plate is equal to  $q/V$ .

Now bring another insulated metal plate *B* near plate *A*. By induction, negative charge is produced on the nearer face of *B* and equal positive charge develops on its farther face (Fig. 18.8 a). The induced negative charge tends to decrease whereas induced positive charge tends to increase the potential of *A*. If plate *B* is earthed (Fig. 18.8 b), the induced positive charge on it, being free, flows to earth. (In reality, it is the negative charge that flows from the earth to the plate. Positive charges in the plate are immobile.) But negative charge will stay as it is bound to positive charge on *A*. Due to this induced negative charge on *B*, the potential of *A* decreases and its capacitance increases.

Hence, we can say that *capacitance of an insulated conductor can be increased by bringing near it an uncharged earthed conductor*.

This is the basic principle of a capacitor. Capacitors are used for storing large amounts of electric charge and hence electrical energy in a small space for a small interval of time.

### A Parallel Plate Capacitor

A parallel plate capacitor is one of the simplest capacitors in which two parallel metallic plates, each of area *A*, are separated from one another by a small distance *d*. An insulating medium like air, paper, mica, glass etc separates the plates. The plates are connected to the terminals of a battery, as shown in Fig. 18.9. Suppose that these plates acquire  $+q$  and  $-q$  charge when the capacitor is fully charged. These charges set up a uniform electric field **E** between the plates. When the separation *d* is small compared to the size of the plates, distortion of electric field at the boundaries of the plates can be neglected.

If  $\sigma$  is surface charge density on either plate, the magnitude of electric field between the plates is given by

$$E = \frac{\sigma}{\epsilon_0} = \frac{q}{\epsilon_0 A}$$

and the potential difference between the plates is given by

$$V = Ed$$

Hence, capacitance of a parallel plate capacitor, whose plates are separated by *d* and have air in-between them is given by

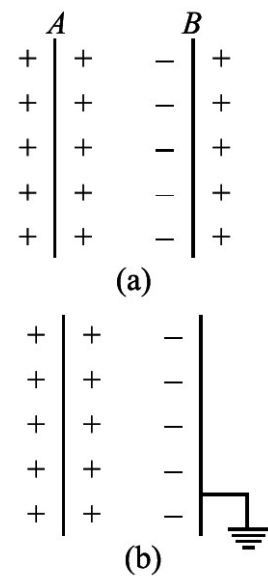


Fig. 18.8 : Working principle of a capacitor

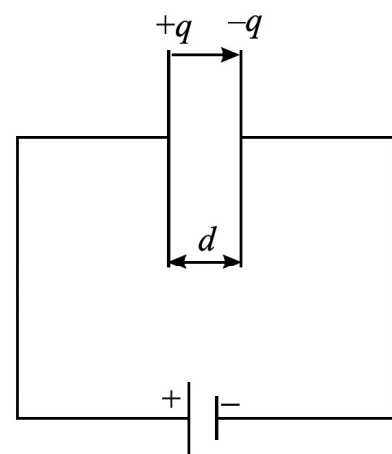


Fig. 18.9 : Working principle of a capacitor

$$C_0 = \frac{q}{V} = \frac{q}{qd/\epsilon_0 A}$$

$$= \frac{\epsilon_0 A}{d} \quad (18.26)$$

It shows that capacitance of a parallel plate capacitor is directly proportional to the area of the plates and inversely proportional to their separation. It means that to obtain high capacitance, area of the plates should be large and separation between them should be small.

If the plates of a capacitor are separated by a dielectric material other than air or vacuum, the capacitance of a parallel plate capacitor is given by

$$C = \frac{\epsilon A}{d} = \frac{k \epsilon_0 A}{d}$$

where  $\epsilon$  is called permittivity of the medium. Therefore, we find that capacitance of a dielectric filled parallel plate capacitor becomes  $K$  times the capacitance with air or vacuum as dielectric :

$$C = KC_0 \quad (18.27)$$

### Relative Permittivity or Dielectric Constant

We can also define dielectric constant by calculating the force between the charges. According to Coulomb's law, the magnitude of force of interaction between two charges  $q_1$  and  $q_2$  separated by a distance  $r$  in vacuum is :

$$F_v = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r^2} \quad (18.28)$$

where  $\epsilon_0$  is the permittivity of free space.

If these charges are held at the same distance in a material medium, the force of interaction between them will be given by

$$F_m = \frac{1}{4\pi\epsilon} \frac{q_1 q_2}{r^2} \quad (18.29)$$

On combining Eqns. (18.28) and (18.29), we get

$$\frac{F_v}{F_m} = \frac{\epsilon}{\epsilon_0} = \epsilon_r \quad (18.30)$$

where  $\epsilon_r$  (or  $K$ ) is relative permittivity. It is also termed as dielectric constant of the medium. Note that it is the ratio of permittivity of the material medium to the permittivity of free space. We can also define the dielectric constant of a medium as the ratio of the electrostatic force of interaction between two point charges held at certain distance apart in air or vacuum to the force of interaction between them held at the same distance apart in the material medium.

The dielectric constant can also be expressed as

$$K = \frac{\text{Capacitance with dielectric between the plates}}{\text{Capacitance with vacuum between the plates}}$$

$$= \frac{C_m}{C_0}$$

Thus

$$C_m = KC_0 \quad (18.31)$$

For metals,  $K = \infty$ , for mica  $K \approx 6$ , and for paper  $K = 3.6$ .

## 18.4) GROUPING OF CAPACITORS

Capacitors are very important elements of electrical and electronic circuits. We need capacitors of a variety of capacitances for different purposes. Sometimes a capacitance of a proper value may not be available. In such situations, grouping of capacitors helps us to obtain desired (smaller or larger) value of capacitance with available capacitors. Two most common capacitor groupings are :

- Series grouping, and
- Parallel grouping.

### Parallel Grouping of Capacitors

In parallel grouping, one plate of each capacitor is connected to one terminal and the other plate is connected to another terminal of a battery, as shown in Fig. 18.10. Let  $V$  be the potential difference applied to the combination between points  $A$  and  $B$ . Note that *in parallel combination, potential difference across each capacitor is the same*. Therefore the charge on each capacitor will be different, say  $q_1$ ,  $q_2$  and  $q_3$  such that

$$\begin{aligned} q_1 &= C_1 V \\ q_2 &= C_2 V \\ q_3 &= C_3 V \end{aligned} \quad (18.32)$$

As charge is conserved the total charge on all the capacitors of the combination is:

$$\begin{aligned} q &= q_1 + q_2 + q_3 \\ q &= (C_1 + C_2 + C_3 + \dots) V \end{aligned} \quad (18.33)$$

Let  $C_p$  be the equivalent capacitance in parallel combination. Then

$$q = C_p V$$

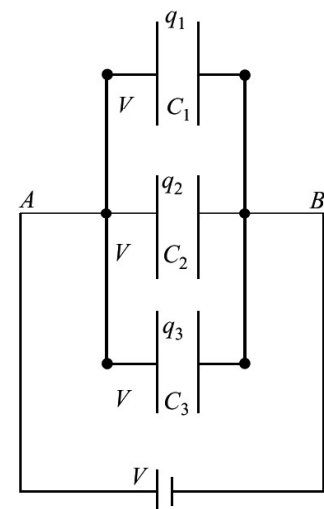


Fig. 18.10 : Capacitors joined in parallel

From these relations, we get

$$q = C_p V = (C_1 + C_2 + C_3)V$$

In general, we can write

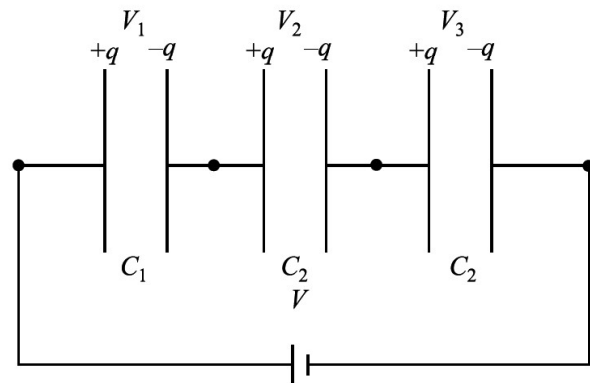
$$C_p = C_1 + C_2 + C_3 = \sum_{i=1}^n C_i \quad (18.34)$$

Thus, we see that *equivalent capacitance of a number of capacitors joined in parallel is equal to the sum of the individual capacitances.*

Remember that in parallel combination, all the capacitors have the same potential *difference between their plates but charge is distributed in proportion to their capacitances.* Such a combination is used for charge accumulation.

### Series Grouping of Capacitors

In the series combination of capacitors, the first plate of the first capacitor is connected to the electrical source. The second plate of the first capacitor is connected to the first plate of the second capacitor. The second plate of second capacitor is connected to first plate of the next capacitor of the combination and so on. The second plate of last capacitor of the combination is connected to the electrical source, as shown in Fig. 18.11. Let  $+q$  unit of charge be given to the first plate of capacitor  $C_1$  from the source. Due to electrical induction, as explained in the principle of capacitor,  $-q$  charge appears on the inner side of right plate of  $C_1$  and  $+q$  charge develops on the outer side of the second plate of  $C_2$ . The  $-q$  unit of charge flows to the first plate of  $C_2$  and so on. Thus, each capacitor receives the same charge of magnitude  $q$ . As their capacitances are different, potential difference across these capacitors will be



**Fig. 18.11** : Capacitors in series grouping. The amount of charge on each capacitor plate is same.

$$V_1 = \frac{q}{C_1}, V_2 = \frac{q}{C_2}, V_3 = \frac{q}{C_3} \quad (18.35)$$

If  $C_s$  is the total capacitance of the series grouping, then

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

and 
$$V = V_1 + V_2 + V_3 \quad (18.36)$$

Hence 
$$\frac{q}{C_s} = \frac{q}{C_1} + \frac{q}{C_2} + \frac{q}{C_3}$$

or 
$$\frac{1}{C_s} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \quad (18.37)$$

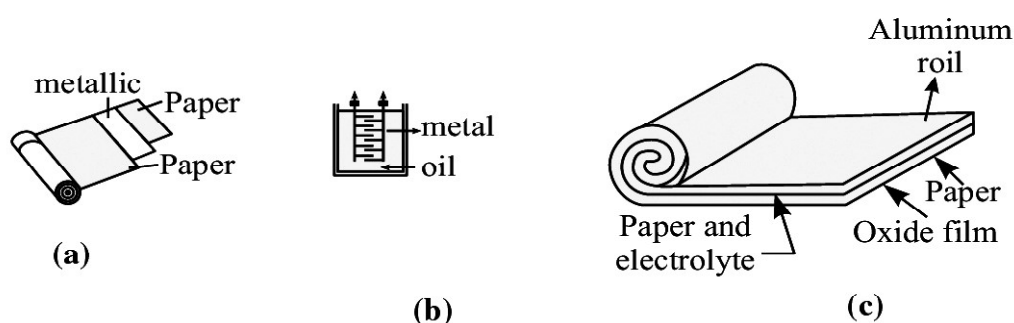
For  $n$  capacitors joined in series, we can write

$$\frac{1}{C_s} = \sum_{i=1}^n \frac{1}{C_i}$$

### Types of Capacitors

There are three common varieties of capacitors in commercial use. Their schematic diagrams are shown in Fig. 18.12.

- Paper capacitor:** Several large thin sheets of paraffin impregnated paper or mylar are cut in proper size (rectangular). Several sheets of metallic foils are also cut to the same size. These are spread one over the other alternately. The outer sheet is mylar, then over it a sheet of metal foil, again over it a sheet of mylar and then a sheet of metal foil and so on. The entire system is then rolled as shown in Fig. 18.12 (a) in the form of a cylinder to form a small device.
- Metal plate capacitors:** A large number of metals are alternately joined to two metal rods as shown in Fig. 18.12 (b). The entire plate system is immersed in silicon oil which works as dielectric material between the plates. High voltage capacitors are usually of this type. Variable capacitors of micro farad capacitance are usually of this type and use air as dielectric. One set of plates is fixed and the other set is movable. The movable plates, when rotated, change their effective area, thereby changing the capacitance of the system. You might see such capacitors in a radio receiver. Variable capacitance helps in tuning to different radio stations.



**Fig. 18.12 :** Different types of capacitors : (a) paper capacitor, (b) variable capacitor, and (c) electrolytic capacitor

- Electrolytic capacitor:** An electrolytic capacitor is shown Fig. 18.12 (c). A metal foil is rolled in the shape of a cylinder with increasing diameter so that there is always a space between one surface and the other. The system is immersed in an electrolyte in the form of a solution. This solution is

conducting because of ions in the solution. A voltage is applied between the electrolyte and the metallic foil. Because of the conducting nature of the electrolyte, a thin layer of metal oxide, which is an insulator, is formed on the foil. The oxide layer works as dielectric material. Since the dielectric layer is extremely thin, the system provides a very high value of capacitance. It is important in this type of capacitor to mark the **positive** and **negative** terminals. A wrong connection of positive and negative terminals removes the oxide layer. (The capacitor then starts conducting.) This type of capacitor is used in storing large amount of charge at low voltage.

Thus, *the reciprocal of equivalent capacitance of any number of capacitors connected in series is equal to the sum of the reciprocals of individual capacitances.* From the above relation, you will agree that  $C_s$  is less than the least of  $C_1$ ,  $C_2$ , and  $C_3$ .

Note that *all the capacitors in series grouping have the same amount of charge but the potential difference between their plates are inversely proportional to their capacitances.* It means that the capacitor with minimum capacitance of the combination will have maximum potential difference between its plates.

### Example 18.3

The capacitance of a parallel plate air capacitor is  $22.0 \mu\text{F}$ . The separation between the plates is  $d$ . A dielectric slab of thickness  $d/2$  is put in-between the plates. Calculate the effective capacitance, if the dielectric constant  $K = 5$ .

#### Solution :

The Capacitance of the air capacitor is given by

$$C_0 = \frac{\epsilon_0 A}{d} = 22.0 \mu\text{F}$$

The new system can be considered as a series combination of two capacitors:

$$C_1 = \frac{K \epsilon_0 A}{d/2} = \frac{2K \epsilon_0 A}{d} = 2KC_0$$

and

$$C_2 = \frac{\epsilon_0 A}{d/2} = \frac{2\epsilon_0 A}{d} = 2C_0$$

The effective capacitance  $C$  is given by

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}$$

or

$$C = \frac{C_1 C_2}{C_1 + C_2}$$

$$= \frac{2KC_0 \times 2C_0}{2KC_0 + 2C_0}$$

$$\begin{aligned}
 &= \frac{2KC_0}{K+1} \\
 &= \frac{10 \times 22 \times 10^{-6} \text{F}}{6} \\
 &= 36.7 \mu\text{f}
 \end{aligned}$$

### Intext Questions 18.2

1. Write the fundamental units of capacitance.
2. What is the potential difference between two points separated by a distance  $d$  in a uniform electric field  $\mathbf{E}$ ?
3. The usual quantities related with an air capacitor are  $C_0$ ,  $E_0$  and  $V_0$ . How are these related with  $C$ ,  $\mathbf{E}$  and  $V$  of the same capacitor filled with dielectric constant  $K$ ?
4. Calculate the area of air filled capacitor plate when the separation between the plates is 50 cm and capacitance is 1.0  $\mu\text{F}$ .

### Energy Stored in a Capacitor

The charging of a capacitor can be visualized as if some external agent, say a battery, pulls electrons from the positive plate of a capacitor and transfers them to the negative plate. Some work is done in transferring this charge, which is stored in the capacitor in the form of electrostatic potential energy. This energy is obtained from the battery (stored as chemical energy). When this capacitor is discharged through a resistor, this energy is released in the form of heat.

Let us assume that an uncharged capacitor, when connected to a battery, develops a maximum charge  $q$ . The charging takes place slowly. The initial potential difference between the capacitor plates is zero and the final potential difference is  $V$ . The average potential difference during the entire process of charging is

$$\frac{0+V}{2} = \frac{V}{2} = \frac{q}{2C}$$

The work done during charging is given by

$$\begin{aligned}
 W &= \text{Charge} \times \text{potential difference} \\
 &= q \frac{q}{2C} = \frac{1}{2} \frac{q^2}{C}
 \end{aligned}$$

Hence potential energy

$$U = \frac{1}{2} qV = \frac{1}{2} \frac{q^2}{C} = \frac{1}{2} CV^2 \quad (18.38)$$



This energy is stored in the electric field between the plates. The stored energy is directly proportional to the capacitance. It also increases as potential difference increases. However, every capacitor can store only a limited amount of energy. An automatic discharge will take place when the potential difference becomes more than its threshold value.

It is dangerous to touch the plates of a charged capacitor. The capacitor may get discharged through your body resulting in an electric shock. Such a shock could be fatal for high value capacitors when fully charged.

## 18.5 DIELECTRICS AND DIELECTRIC POLARIZATION

We know that dielectrics are insulating materials, which transmit electric effects without conducting. Dielectrics are of two types : **non-polar** and **polar**. We now learn about these.

- (a) **Non-polar dielectrics** : In the molecules of non-polar dielectrics, the centre of positive charge coincides with the centre of negative charge. Each molecule has zero dipole moment in its normal state. These molecules are mostly symmetrical such as nitrogen, oxygen, benzene, methane,  $\text{CO}_2$ , etc.
- (b) **Polar dielectrics** : Polar dielectrics have asymmetric shape of the molecules such as water,  $\text{NH}_3$ ,  $\text{HCl}$  etc. In such molecules, the centres of positive and negative charges are separated through a definite distance and have finite permanent dipole moment.

When a non-polar dielectric is held in an external electric field, the centre of positive charge in each molecule is pushed in the direction of  $\mathbf{E}$  and the centre of negative charge is displaced in the direction opposite to  $\mathbf{E}$ . Because of external electric field, centres of positive and negative charges in the non-polar dielectric molecules are separated. Dielectric is then said to be polarized and a tiny dipole moment develops in each molecule. In fact, the force due to external electric field pulling the charge centres apart balances the force of mutual attraction between the centres (i.e. equilibrium is set) and the molecule is said to be polarised. Induced dipole moment  $\mathbf{p}$  acquired by the molecule may be written as

$$\mathbf{p} = \alpha \epsilon_0 \mathbf{E}$$

where  $\alpha$  is constant of proportionality and is called atomic/molecular polarizability.

Let us now consider a non-polar slab ABCD placed in an electric field  $\mathbf{E}$  maintained between the plates of a capacitor. As shown in Fig. 18.13, the dielectric slab gets polarised. The nuclei of dielectric molecules are displaced towards the negative plate and electrons towards the positive plate. Because of polarisation, an electric field  $\mathbf{E}_p$  is produced within the dielectric, which is opposite to  $\mathbf{E}$ . Hence, due to the presence of a non-polar dielectric, the field between the plates is reduced, i.e. effective electric field in a

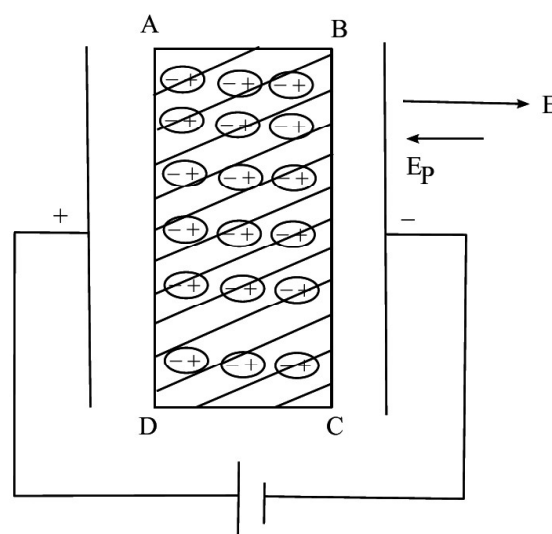


Fig. 18.13 : A dielectric slab between the charged capacitor plates.

polarised dielectric is given by

$$\mathbf{E} \text{ (effective)} = \mathbf{E} - \mathbf{E}_p \quad (18.39)$$

Thus, the potential difference between the capacitor plates is correspondingly reduced (as  $V = Ed$ ), increasing the value of capacitance of the capacitor (as  $C = q/V$ ).

### Application of Electrostatics

Electrostatics provides basis for the theory of electromagnetics, apart from useful assistance in many fields of science and technology.

- Capacitors are essential parts of most electronic and electrical circuitry. These play a very crucial role in power transmission.
- Gold leaf electroscope – the simple device used for detecting charge, paved the way for cosmic ray research.
- Lightning conductor devised by Benjamin Franklin is still used to protect skyscrapers from the strokes of lightning and thunder.
- The working of photocopiers, so common these days, is based on the principle of electrostatics.

### Intext Questions 18.3

- Two capacitors  $C_1 = 12 \text{ mF}$  and  $C_2 = 4 \text{ mF}$  are in group connections. Calculate the effective capacitance of the system when they are connected (a) in series (b) in parallel.
- Four capacitors are connected together as shown in Fig. 18.14. Calculate the equivalent capacitance of the system.
- An air capacitor  $C = 8 \text{ mF}$  is connected to a 12V battery. Calculate
  - the value of  $Q$  when it is fully charged?
  - the charge on the plates, when slab of dielectric constant  $K = 5$  fills the gap between the plates completely.
  - potential difference between the plates; and
  - capacitance of the new capacitor
- A parallel plate capacitor of capacitance  $C_0$  is connected to a battery and charged to a potential difference  $V_0$ . After disconnecting the battery, the gap between the plates is completely filled with a slab of dielectric constant  $K$ . How much energy is stored in the capacitor (a) in the first state? (b) in the second state? and (c) which one is larger and why?

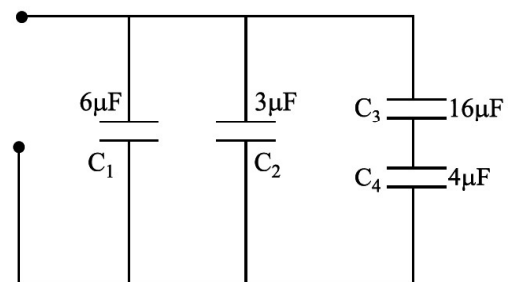


Fig. 18.14 : Grouping of capacitors

## WHAT YOU HAVE LEARNT

- The potential at any point in an electric field is equal to the work done against the electric field in moving a unit charge from infinity to that point.
- Work done in transferring a charge from one point to another in an electrostatic field is path independent.
- If one joule of work is done in bringing a test charge of one coulomb from infinity to a point in the field, we say that potential at that point is one volt.
- Electric potential due to a dipole is zero at every point on the equatorial line of the dipole.
- In an equipotential surface, every point has same electric potential.
- At any point the electric field is equal to the ratio of negative change of potential with distance (called potential gradient) gives the field.
- Electrostatic shielding is the phenomenon of protecting a region of space from electric field.
- Capacitance of a conductor depends on its shape, size and nature of medium, rather than its material.
- The capacitance of a dielectric filled parallel plate capacitor becomes K times the capacitance with air or vacuum as dielectric.
- Relative permittivity is the ratio of capacitance with dielectric between the plates to the capacitance with air or vacuum between the plates.
- In series combination of capacitors, the equivalent capacitance is less than the least of any of the individual capacitances.
- In parallel combination of capacitors, the equivalent capacitance is equal to the sum of individual capacitances.
- Due to the presence of a non-polar dielectric, the field between the plates of a capacitor is reduced.

## TERMINAL EXERCISE

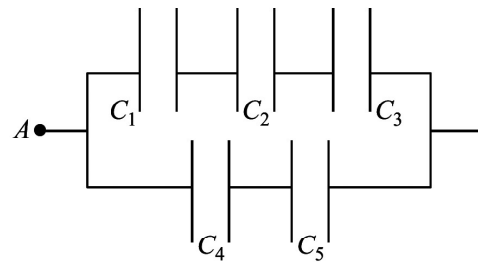
1. Calculate the potential at a point P at a distance of 30 cm from a point charge  $q = 20 \mu\text{C}$
2. Three point charges  $q_1$ ,  $q_2$  and  $q_3$ , each of magnitude  $200 \mu\text{C}$ , are placed at the corners A, B and C respectively of an equilateral triangle. The length of the side is 10cm. Calculate the potential energy of the system.
3. The potential difference between the plates of a capacitor separated by 3mm is 12.0 V. Calculate the magnitude of E between the plates?
4. Two ions having charges  $+e$  and  $-e$  are  $4.0 \times 10^{-10}$  m apart. Calculate the potential energy of the system.

- The plates A and B of a parallel plate capacitor have a potential difference of 15 V. A proton ( $m = 1.67 \times 10^{-27}$  kg) is moved from the positive plate A to B. Calculate the speed of the proton near plate B.
- Show that dimensionally the quantities  $Vq$  and  $(\frac{1}{2})mv^2$  are equivalent. The symbols carry the usual meaning.
- Under what condition, the electric field between the plates of a parallel plate capacitor is uniform?
- A metallic sphere of radius  $r$  has a charge  $+q$ . Calculate the work done in moving a test charge  $q_0$  from one end of a diameter to its other end.
- A parallel plate air capacitor of value  $C_0$  is charged to a potential  $V_0$  between the plates and  $+q_0$  is charge on one plate. Separation between plates is  $d$ . A dielectric of dielectric constant  $K = 3$  fills the space between the plates. Which of these quantities will change and why. (i) capacitance (ii) charge (iii) potential difference and (iv) field density?
- Examine the following network of capacitors. The potential difference between A and B is 16V :

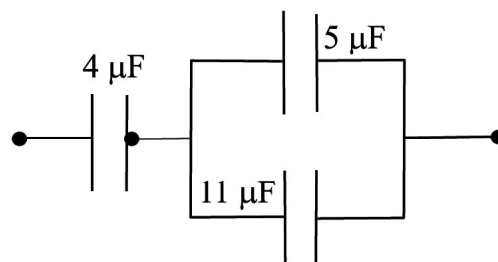
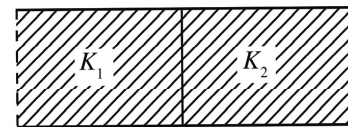
$$C_1 = 2 \mu\text{F}, C_2 = 4 \mu\text{F}, C_3 = 8 \mu\text{F},$$

$$C_4 = 3 \mu\text{F}, C_5 = 3 \mu\text{F}$$

Calculate (a) the effective capacitance between A and B, (b) the charge on each capacitor, and (c) the potential difference across each capacitor.

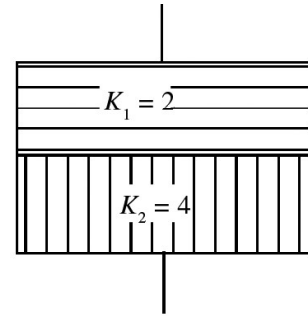


- The value of capacitance of an air capacitor is  $8 \mu\text{F}$ . Two dielectrics of identical size fill the space between the plates as shown. Dielectric constants are  $K_1 = 3.0$  and  $K_2 = 6.0$ . Calculate the value of the new capacitance.
- Calculate the equivalent capacitance of the following system.



- A  $3.0 \mu\text{F}$  air capacitor is charged to a potential 12.0 V. A slab of dielectric constant  $K = 7$  is made to fill the space. Calculate the ratio of the energies stored in the two systems.
- A dipole of dipole moment  $P = 3.5 \times 10^{-15}$  Cm is placed in a uniform electric field  $E = 2.0 \times 10^4$  NC<sup>-1</sup>. The dipole makes an angle of  $60^\circ$  with the field. Calculate the (a) Potential energy of the dipole and (b) the torque on the dipole.

15. The capacitance of a parallel plate air capacitor is  $12 \mu\text{F}$ . The separation between the plates is  $8 \text{ mm}$ . Two dielectric slabs of the same size fill the air space. Calculate the new value of capacitance.



**ANSWERS TO INTEXT QUESTIONS**

**18.1**

1. The potential at  $r$  ( $r > R$ )

$$V = \frac{1}{4\pi\epsilon_0} \frac{q}{r}$$

2. The field around a point charge possesses spherical symmetry. Thus every point on the surface of the sphere is equipotential. And no work is done when a charge moves on an equipotential surface

3.  $E = -\frac{dV}{dr}$  Since  $V$  is constant,  $E$  is zero.

We can obtain the same result using Eqn. (18.22) :

$$E = \frac{V_A - V_B}{d}. \text{ Since } V_A = V_B, E \text{ is zero}$$

4. No. Not necessarily. When  $E = 0$ , the potential is either constant or zero.
5. Two equipotential surfaces never intersect. If they do so, at the point of intersection we can draw two normals giving two directions of electric field, which is not possible.

**18.2**

$$1. \quad C = \frac{Q}{V} = \frac{\frac{Q}{\text{Work done}}}{\text{Charge}} = \frac{Q \times Q}{\text{Work done}}$$

$$= \frac{Q^2}{\text{N.m.}}$$

The basic unit is

$$A = \frac{C}{s}$$

$$\therefore C^2 = A^2 s^2 \text{ and newton} = \text{mass} \times \text{acc} = \text{kg} \frac{\text{m}}{\text{s}^2}$$

$$\begin{aligned} \text{Capacitance} &= \frac{C^2}{\text{Nm}} = \frac{A^2 s^2}{\text{kg} \frac{\text{m}^2}{\text{s}^2}} = \frac{A^2 s^4}{\text{kg m}^2} \\ &= A^2 s^4 (\text{kg m}^2)^{-1} \end{aligned}$$

2. In a capacitor,  $E$  is uniform between these plates. Potential difference between the plates

$$V_A - V_B = E \times d.$$

3.  $C_0, E_0, V_0$  for air capacitor and  
 $C, E, V$  for dielectric capacitor. Then

$$k = \frac{C}{C_0}, k = \frac{V_0}{V} \quad k = \frac{E_0}{E}.$$

4.  $C = 1.0 \mu\text{F} = 1.0 \times 10^{-6} \text{F}$ .  
 $d = 50 \text{cm} = 0.5 \text{m}$ .

$$C = \frac{\epsilon_0 A}{d}$$

$$\therefore A = \frac{Cd}{\epsilon_0}. \text{ Since } \epsilon_0 = 8.85 \times 10^{-12},$$

$$\begin{aligned} A &= \frac{1.0 \times 10^{-6} \times 0.5}{8.85 \times 10^{-12}} \\ &= \frac{5 \times 10^{-7}}{8.85 \times 10^{-12}} \\ &= 0.56 \times 10^5 \text{ m}^2 \end{aligned}$$

### 18.3

- (a) 3 mF      (b) 16 mF
- 12.2  $\mu\text{F}$
- (a) 96 mC      (b) 96 mC      (c) 2.4 V      (d) 40 mF
- (a)  $\frac{1}{2} C_0 V_0^2$       (b)  $\frac{1}{2} \frac{(C_0 V_0)^2}{C_0 R} = \frac{1}{2k} C_0 V_0^2$

- (c) The energy in the first case is more, because same energy is used up for sucking in the dielectric slab.

**ANSWERS TO TERMINAL EXERCISE**

1.  $6 \times 10^5 \text{ V}$
2.  $1.08 \times 10^4 \text{ J}$
3.  $4 \times 10^3 \text{ Vm}^{-1}$
4.  $-5.76 \times 10^{-19} \text{ J}$
5.  $5.36 \times 10^4 \text{ m/s}$
  
10. (a)  $\frac{37}{14} \mu\text{F}$ ,  
(b)  $\frac{128}{7} \mu\text{C}$ ,  $\frac{128}{7} \mu\text{C}$ ,  $\frac{128}{7} \mu\text{C}$ ,  $24 \mu\text{C}$ ,  $24 \mu\text{C}$   
(c)  $\frac{64}{7} \text{V}$ ,  $\frac{32}{7} \text{V}$ ,  $\frac{16}{7} \text{V}$ ,  $8 \text{V}$ ,  $8 \text{V}$
11.  $36 \mu\text{F}$ .
12.  $\frac{16}{5} \mu\text{F}$
13.  $1 : 7$
14. (a)  $3.5 \times 10^{-11} \text{ J}$     (b)  $6 \times 10^{-11} \text{ Nm}$ .
15.  $32 \mu\text{F}$



# CURRENT ELECTRICITY

## INTRODUCTION

In our daily life we use electricity for various activities. The electric lamps and tubes light our houses, we listen music on a tape recorder or radio, see different programmes on television, enjoy cool breeze from electric fan or cooler, and use electric pump to irrigate fields. In fact, electricity is a unique gift of science to mankind. We can not imagine life without electricity in the modern world. At home you might have observed that as soon as switch on an electric lamp, it starts glowing. Why does it happen? What is the function of a switch?

In the preceding lessons you have studied about static electric charges and forces between them. In this lesson, you will learn about electric charges in motion. You will also learn that the rate of flow of charge through a conductor depends on the potential difference across it. You will also study the distribution of current in circuits and Kirchhoff's laws which govern it.

Physics is an experimental science and the progress it has made to unfold laws of nature became possible due to our ability to verify theoretical predictions or reproduce experimental results. This has led to continuous improvement in equipment and techniques. In this lesson you will learn about potentiometer, which is a very versatile instrument. It can be used to measure resistance as well as electro-motive force using null method.

## OBJECTIVES

After studying this lesson, you should be able to

- *state Ohm's law and distinguish between ohmic and non-ohmic conductors;*
- *obtain equivalent resistance for a series and parallel combination of resistors;*
- *apply Kirchhoff's rules to closed electrical circuits;*
- *apply Wheatstone bridge equation to determine an unknown resistance;*
- *explain the principle of potentiometer and apply it to measure the e.m.f and internal resistance of a cell.*

### Free and Bound Electrons

An atom is electrically neutral, i.e. as many negatively charged electrons revolve around the nucleus in closed orbits as there are positively charged protons inside it. The electrons are bound with the nucleus through Coulomb (attractive) forces.



Farther the electrons from the nucleus, weaker is the Coulomb force. The electrons in the outermost orbit are, therefore, most loosely bound with the nucleus. These are called *valence electrons*. In metallic solids, the valence electrons become free to move when a small potential difference is applied.

## 19.1) ELECTRIC CURRENT

You have studied in the previous lesson that when a potential difference is applied across a conductor, an electric field is set up within it. The free electrons move in a direction opposite to the field through the conductor. This constitutes an electric current. Conventionally, *the direction of current is taken as the direction in which a positive charge moves*. The electrons move in the opposite direction. To define current precisely, let us assume that the charges are moving perpendicular to a surface of area  $A$ , as shown in Fig. 19.1. The current is the rate of flow of charge through a surface area placed perpendicular to the direction of flow. If charge  $\Delta q$  flows in time  $\Delta t$ , the average current is defined as :

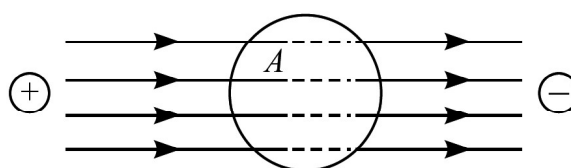


Fig. 19.1 : Motion of charges inside a conductor of surface area  $A$

$$I_{av} = \frac{\Delta q}{\Delta t} \quad (19.1)$$

If the rate of flow of charge varies with time, the current also varies with time. The instantaneous current is expressed as :

$$I = \frac{dq}{dt} \quad (19.2)$$

**The electric current through a conductor is the rate of transfer of charge across a surface placed normal to the direction of flow.** Another two important points to be noted for flow of current are:

1. For current flow the circuit should be closed. In open circuit current will not flow.
2. For current flow there should exist a potential difference between the two points. If both the points are at the same potential then current will not flow. For any electrical network this basic rules apply

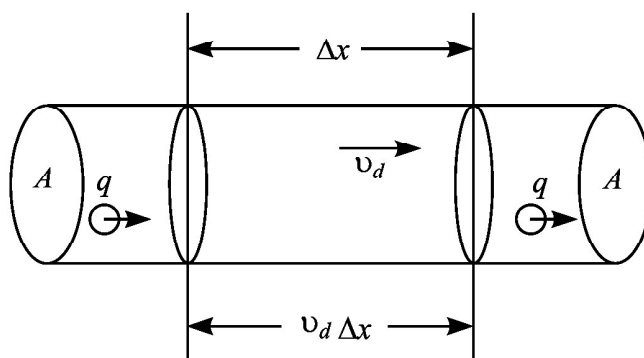
The SI unit of current is ampere. Its symbol is  $A$  :

$$1 \text{ ampere} = \frac{1 \text{ coulomb}}{1 \text{ second}} \quad (19.3)$$

The smaller units of current are milliampere,  $1 \text{ mA} = 10^{-3} \text{ A}$ , and microampere,  $1 \mu\text{A} = 10^{-6} \text{ A}$ . The current can arise due to flow of negative charges (electrons), as in metals. In a semiconductor, flow of electrons (negative charge) and holes constitutes current.

Holes are vacancies in a crystal. These are taken as positively charged particles having the same amount of charge as that on an electron.

Let us consider a conductor of cross sectional area  $A$  shown in Fig. 19.2. The volume element for a length  $\Delta x$  is  $A \Delta x$ . If  $n$  is the number of electrons per unit volume, the number of electrons in this volume element will be  $nA\Delta x$ . The total charge in this volume element is  $\Delta q = nA\Delta xe$ , where  $e$  is charge of the electron. If electrons drift with a speed  $v_d$  due to thermal energy, the distance travelled in time  $\Delta t$  is  $\Delta x = v_d \Delta t$ . On substituting this value of  $\Delta x$  in the expression for  $\Delta q$ , we find that total charge in the volume element under consideration is given by



**Fig. 19.2 :** The charges move with a speed  $v_d$  through a surface of area  $A$ . The number of charges in a length  $\Delta x$  is  $nA v_d \Delta t$ .

$$\Delta q = nAe v_d \Delta t$$

so that

$$\frac{\Delta q}{\Delta t} = I = nAe v_d \quad (19.4)$$

You will learn more about the drift velocity in sec.19.9.

## 19.2) OHM'S LAW

In 1828, Ohm studied the relation between current in a conductor and potential difference applied across it. He expressed this relation in the form of a law, known as **Ohm's law**.

### George Simon Ohm (1787-1854)

German physicist, George Simon Ohm is famous for the law named after him. He arrived at the law by considering an analogy between thermal and electrical conduction. He also contributed to theory of sirens, interference of polarised light in crystals etc. Ohm, the practical unit of resistance, is named in his honour.



According to Ohm's law, **the electric current through a conductor is directly proportional to the potential difference across it, provided the physical conditions such as temperature and pressure remain unchanged.**

Let  $V$  be the potential difference applied across a conductor and  $I$  be the current flowing through it. According to Ohm's law,

$$V \propto I$$

or

$$V = RI$$

$$\Rightarrow \quad \frac{V}{I} = R \quad (19.5)$$

where constant of proportionality  $R$  signifies the electrical resistance offered by a conductor to the flow of electric current. *Resistance is the property of a conductor by virtue of which it opposes the flow of current through it.* The  $I$ - $V$  graph for a metallic conductor is a straight line [Fig. 19.3 (a)].

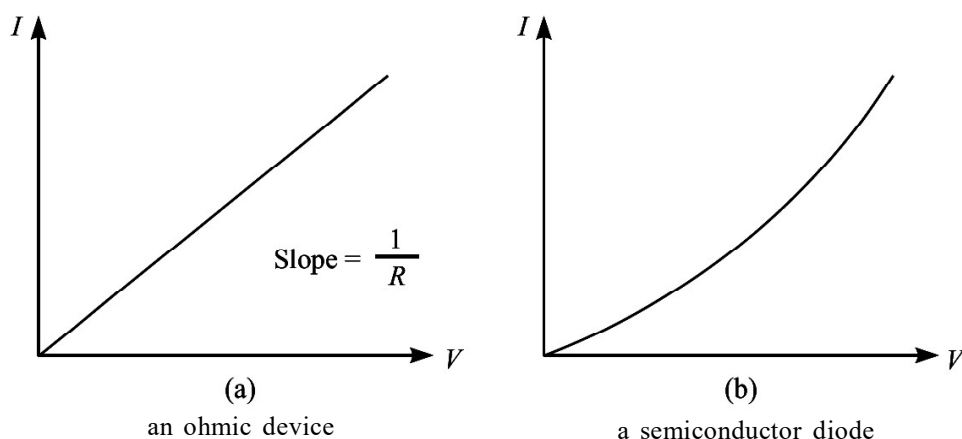


Fig. 19.3 : Current-voltage graph

The SI unit of resistance is ohm. It is expressed by symbol  $\Omega$  (read as omega)

$$1 \text{ ohm} = 1 \text{ volt/1 ampere}$$

Most of the metals obey Ohm's law and the relation between voltage and current is linear. Such resistors are called *ohmic*. Resistors which do not obey Ohm's law are called *non-ohmic*. Devices such as vacuum diode, semiconductor diode, transistors show non ohmic character. For semiconductor diode, Ohm's law does not hold good even for low values of voltage. Fig. 19.3 (b) shows a non-linear  $I$ - $V$  graph for a semiconductor diode.

### Activity 19.1

**Aim :** To study conduction of electricity through an electrolyte.

**Material Required** Ammeter, Voltmeter, a jar containing copper sulphate solution, two copper plates, a battery, plug key, connecting wires and a rheostat.

#### How to Proceed :

1. Set up the apparatus as shown in Fig. 19.4.
2. Plug in the key and note ammeter and voltmeter readings.

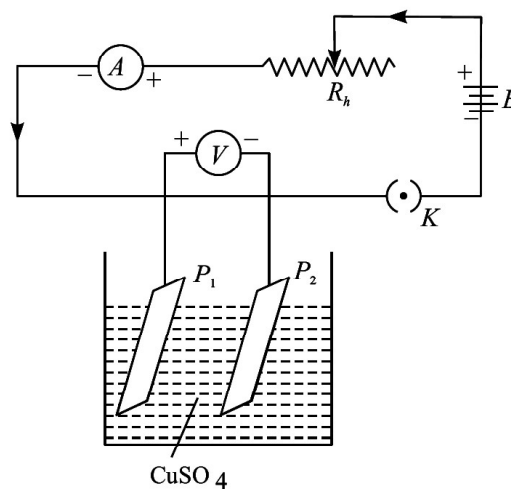


Fig. 19.4 : Electrical conduction through an electrolyte

3. Change the value of ammeter reading by moving the sliding contact of rheostat and note voltmeter reading.
4. Repeat step-3 at least five times and record ammeter and voltmeter readings each time.
5. Repeat the experiment by changing (a) separation between  $P_1$  and  $P_2$ , (b) plate area immersed in electrolyte, and (c) concentration of electrolyte.
6. Plot I-V graph in each case.

### What do you conclude?

- If I-V graph is a straight line passing through the origin, as shown in Fig. 19.5, we say that ionic solution behaves as an ohmic resistor.
- The slope of the graph changes steeply with change in volume of electrolyte between the plates. It means that resistivity of an electrolyte depends not only on its nature but also on the area of the electrodes and the separation between them.

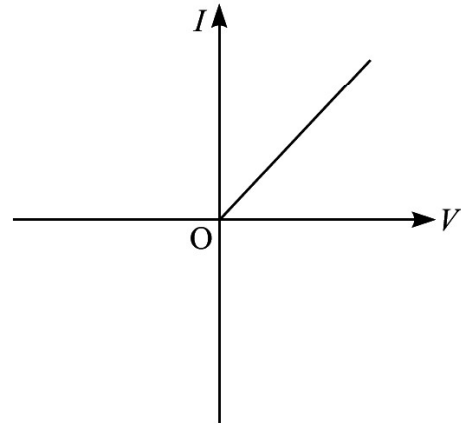


Fig. 19.5 : I-V graph for an ionic solution

### Resistance and Resistivity

Let us now study the factors which affect the resistance of a conductor. You can perform two simple experiments. To do so, set up a circuit as shown in Fig. 19.6.

#### Activity 19.2

Take a long conducting wire of uniform cross section. Cut out pieces of different lengths, say  $l_1, l_2, l_3$ , etc from it. This makes sure that wires have same area of cross-section. Connect  $l_1$  between A and B and note down the current through this wire. Let this current be  $I$ . Perform the same experiment with wires of lengths  $l_2$  and  $l_3$  one by one. Let the currents in the wires be  $I_2$  and  $I_3$  respectively. Plot a graph between  $l^{-1}$  and  $I$ . You will find that the graph is a straight line and longer wires allow smaller currents to flow. That is, longer wires offer greater resistance [Fig. 19.7 (a)]. Mathematically, we express this fact as

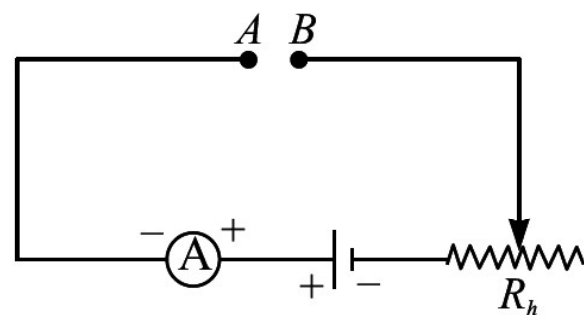


Fig. 19.6 : Electrical circuit to study factors affecting resistance of conductors

$$R \propto l \quad (19.6)$$

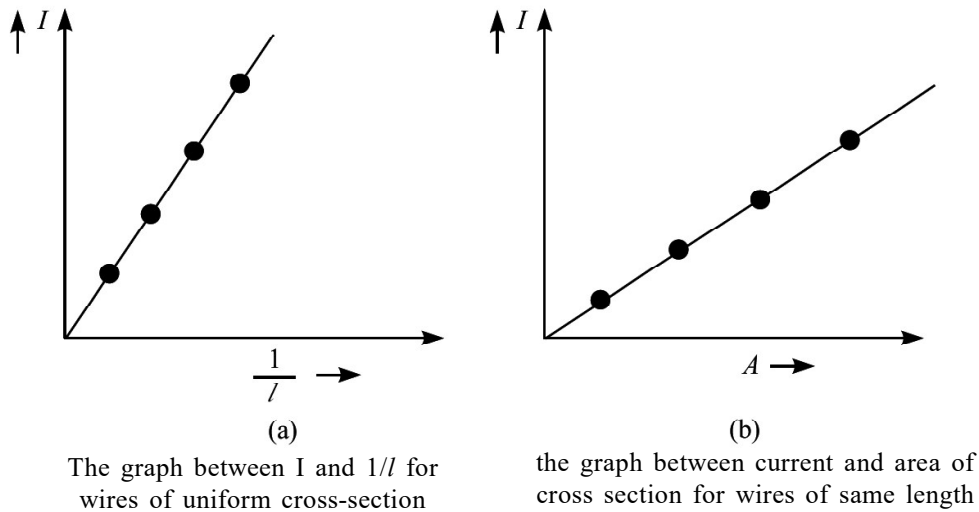


Fig. 19.7

### Activity 19.3

Take wires of the same length of a given material but having different areas of cross section, say  $A_1, A_2, A_3$  etc. Connect the wires between A and B one by one and note down the currents  $I_1, I_2, I_3$  etc. in each case. A plot of  $I$  and  $A$  will give a straight line. Wires of greater cross sectional area allow greater currents to flow. You may say that wires of larger area of cross-section offer smaller resistance [Fig. 19.7 (b)]. Mathematically, we can write

$$R \propto \frac{1}{A} \quad (19.7)$$

On combining Eqns. (19.6) and (19.7), we can write

$$R \propto \frac{l}{A}$$

or

$$R = \rho \frac{l}{A} \quad (19.8)$$

where  $\rho$  is a constant for the material at constant temperature. It is called the **specific resistance** or **resistivity** of the material. By rearranging terms, we can write

$$\rho = \frac{RA}{l} \quad (19.9)$$

If  $l = 1$  m and  $A = 1$  m<sup>2</sup>, then  $\rho = R$  ohm-metre. Thus *resistivity of a material is the resistance offered by a wire of length one metre and area of cross section one m<sup>2</sup>*. The unit of resistivity is ohm metre ( $\Omega\text{m}$ )

Reciprocal of resistivity is called conductivity (specific conductance) and is denoted by  $\sigma$  :

$$\sigma = \frac{1}{\rho} \quad (19.10)$$

Unit of **conductivity** is  $\text{Ohm}^{-1} \text{metre}^{-1}$  or  $\text{mho-metre}^{-1}$  or  $\text{Sm}^{-1}$ .

Resistivity depends on the nature of the material rather than its dimensions, whereas the resistance of a conductor depends on its dimensions as well as on the nature of its material.

You should now study the following examples carefully.

### Example 19.1

In our homes, the electricity is supplied at 220 V. Calculate the resistance of the bulb if the current drawn by it is 0.2 A.

**Solution :**

$$R = \frac{V}{I} = \frac{220 \text{ volt}}{0.2 \text{ amp.}} = 1100 \Omega$$

### Example 19.2

A total of  $6.0 \times 10^{16}$  electrons pass through any cross section of a conducting wire per second. Determine the value of current in the wire.

**Solution :**

Total charge passing through the cross-section in one second is

$$\Delta Q = ne = 6.0 \times 10^{16} \times 1.6 \times 10^{-19} \text{ C} = 9.6 \times 10^{-3} \text{ C}$$

$$\begin{aligned} \therefore I &= \frac{\Delta Q}{\Delta t} = \frac{9.6 \times 10^{-3} \text{ C}}{1 \text{ s}} \\ &= 9.6 \times 10^{-3} \text{ A} \\ &= 9.6 \text{ mA} \end{aligned}$$

### Example 19.3

Two copper wires A and B have the same length. The diameter of A is twice that of B. Compare their resistances.

**Solution :**

From Eqn. (19.8) we know that

$$R_A = \rho \frac{l}{\pi r_A^2} \quad \text{and} \quad R_B = \rho \frac{l}{\pi r_B^2}$$

$$\therefore \frac{R_A}{R_B} = \frac{r_B^2}{r_A^2}$$

Since diameter of A = 2 × diameter of B, we have  $r_A = 2r_B$ . Hence Resistance of B will be four times the resistance of A.

**Example 19.4**

The length of a conducting wire is 60.0 m and its radius is 0.5 cm. A potential difference of 5.0 V produces a current of 2.5 A in the wire. Calculate the resistivity of the material of the wire.

**Solution :**

$$R = \frac{V}{I} = \frac{5.0 \text{ V}}{2.5 \text{ A}} = 2.0 \Omega$$

Radius of the wire = 0.5 cm =  $5.0 \times 10^{-3}$  m

Area of cross section  $A = \pi R^2 = 3.14 \times (5.0 \times 10^{-3})^2 \text{ m}^2 = 78.5 \times 10^{-6} \text{ m}^2$

$$\therefore \rho = \frac{2.0 \times 78.5 \times 10^{-6} \Omega \text{m}^2}{60.0 \text{ m}} = 2.6 \times 10^{-6} \Omega \text{m}$$

**Intext Questions 19.1**

- A current  $I$  is established in a copper wire of length  $l$ . If the length of the wire is doubled, calculate the current due to the same cell.
  - What happens to current in an identical copper wire if the area of cross section is decreased to half of the original value?
- The resistivity of a wire of length  $l$  and area of cross section  $A$  is  $2 \times 10^{-8} \Omega \text{m}$ . What will be the resistivity of the same metallic wire of length  $2l$  and area of cross section  $2A$ ?
- A potential difference of 8 V is applied across the ends of a conducting wire of length 3 m and area of cross section  $2 \text{ cm}^2$ . The resulting current in the wire is 0.15 A. Calculate the resistance and the resistivity of the wire.
- Do all conductors obey Ohm's law? Give examples to support your answer.
- $5 \times 10^{17}$  electrons pass through a cross-section of a conducting wire per second from left to right. Determine the value and direction of current.

**19.3) GROUPING OF RESISTORS**

An electrical circuit consists of several components and devices connected together. Some of these are batteries, resistors, capacitors, inductors, diodes, transistors etc. (They are known as circuit elements.) These are classified as resistive and reactive. The most common resistive components are resistors, keys, rheostats, resistance coils, resistance boxes and connecting wires. The reactive components include capacitors, inductors and transformers. In addition to many other functions performed by these elements individually or collectively,

they control the current in the circuit. In the preceding lesson you learnt how grouping of capacitors can be used for controlling charge and voltage. Let us now discuss the role of combination of resistors in controlling current and voltage.

Two types of groupings of resistors are in common use. These are : **series grouping and parallel grouping**. We define equivalent resistance of the combination as a single resistance which allows the same current to flow as the given combination when the same potential difference is applied across it.

### Series Combination

You may connect many resistors in series by joining them end-to-end such that the same current passes through all the resistors. In Fig. 19.8, two resistors of resistances  $R_1$  and  $R_2$  are connected in series. The combination is connected to a battery at the ends A and D. Suppose that current  $I$  flows through the series combination when it is connected to a battery of voltage  $V$ . Potential differences  $V_1$  and  $V_2$  develop across  $R_1$  and  $R_2$ , respectively. Then  $V_1 = IR_1$  and  $V_2 = IR_2$ . But sum of  $V_1$  and  $V_2$  is equal to  $V$ , i.e.

$$\Rightarrow V = V_1 + V_2 = I R_1 + I R_2$$

If equivalent resistance of this series combination is  $R$ , then

$$V = IR = I (R_1 + R_2)$$

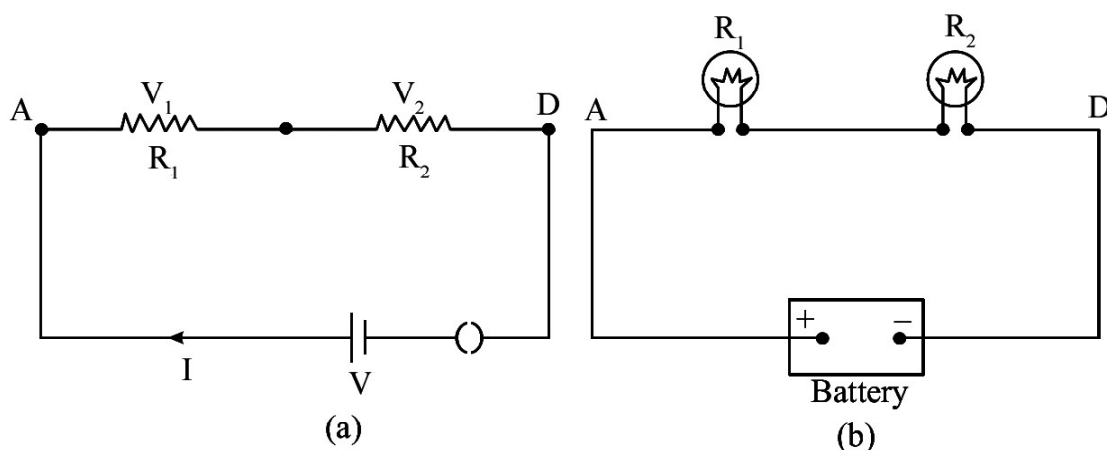
so that

$$R = R_1 + R_2$$

This arrangement may be extended for any number of resistors to obtain

$$R = R_1 + R_2 + R_3 + R_4 + \dots \quad (19.11)$$

That is, the equivalent resistance of a series combination of resistors is equal to the sum of individual resistances. If  $n$  equal resistances are connected in series, then the equivalent resistance is equal to  $nR$  and potential difference across each resistance is  $v/n$ . If we wish to apply a voltage across a resistor (say electric lamp) less than that provided by the constant voltage supply source, we should connect another resistor (lamp) in series with it.



Two resistors connected in series to a battery

two lamps joined in series connected to a dc source.

Fig. 19.8



### Voltage Division in Series Combination of two resistors

We know that  $V = V_1 + V_2$  and

$$V = IR_1 + IR_2 = (R_1 + R_2)I$$

$$V = (R_1 + R_2)I$$

$$I = \frac{V}{R_1 + R_2}$$

Voltage or potential across  $R_1$ , is given by

$$V_1 = IR_1 \text{ but } I = \frac{V}{R_1 + R_2}$$

$$\therefore V_1 = \left( \frac{V}{R_1 + R_2} \right) R_1 \Rightarrow V_1 = \left( \frac{R_1}{R_1 + R_2} \right) V$$

Similarly potential or voltage across  $R_2$  is given by

$$V_2 = IR_2 \text{ but } I = \frac{V}{R_1 + R_2}$$

$$V_2 = \left( \frac{V}{R_1 + R_2} \right) R_2$$

$$V_2 = \frac{R_2}{(R_1 + R_2)} V$$

In general to calculate the potential drop across any resistance it is equal to that resistance divided by total resistance and multiplied by total voltage.

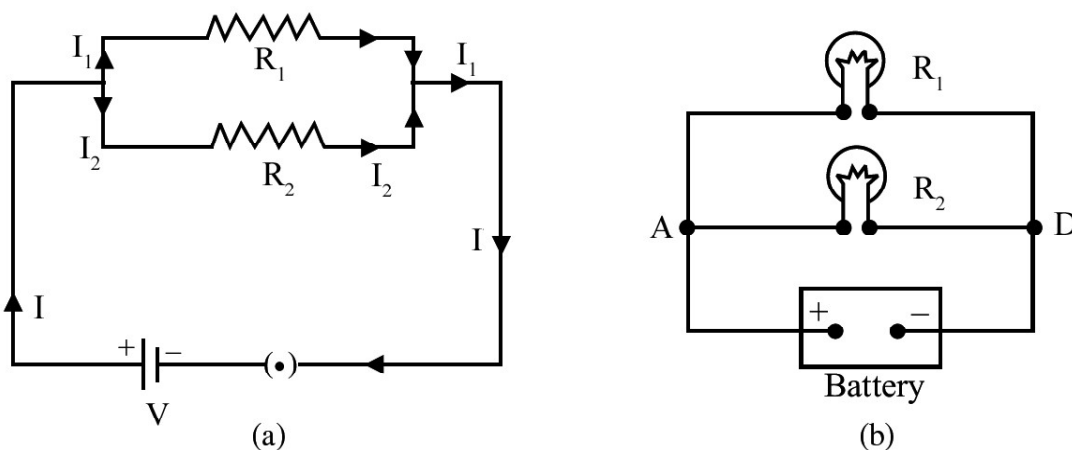
### Parallel Combination

You may connect the resistors in parallel by joining their one end at one point and the other ends at another point. In parallel combination, **same potential difference exists across all resistors**. Fig. 19.9 shows a parallel combination of two resistors  $R_1$  and  $R_2$ . Let the combination be connected to a battery of voltage  $V$  and draw a current  $I$  from the source.

The main current divides into two parts. Let  $I_1$  and  $I_2$  be the currents flowing through resistors  $R_1$  and  $R_2$ , respectively. Then  $I_1 = V/R_1$  and  $I_2 = V/R_2$ .

The main current is the sum of  $I_1$  and  $I_2$ . Therefore, we can write

$$\Rightarrow I = I_1 + I_2 = \frac{V}{R_1} + \frac{V}{R_2}$$



Two resistors connected in parallel. The battery supplies the same voltage to both resistors

lamps connected in parallel to a battery

Fig. 19.9

If the equivalent resistance of combination is  $R$ , we write  $V = IR$  or  $I = V/R$ :

$$I = \frac{V}{R} = \frac{V}{R_1} + \frac{V}{R_2}$$

$$\therefore \frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} \quad (19.12 \text{ a})$$

or

$$R = \frac{R_1 R_2}{R_1 + R_2} \quad (19.12 \text{ b})$$

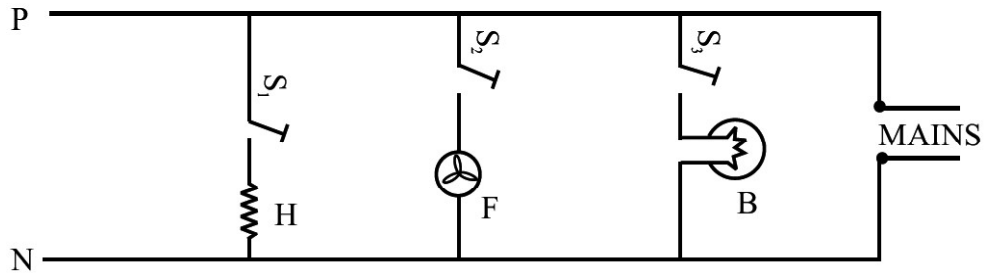
From Eqn. (19.12 a) we note that **reciprocal of equivalent resistance of parallel combination is equal to the sum of the reciprocals of individual resistances**. The process may be extended for any number of resistors, so that

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \frac{1}{R_4} + \dots \quad (19.13)$$

Note that the equivalent resistance of parallel combination is smaller than the smallest individual resistance. If  $n$  equal resistance of value  $R$  are connected in parallel then the equivalent resistance is equal to  $R/n$ . You may easily see this fact by a simple electrical circuit having a resistor of  $2 \Omega$  connected across a  $2 \text{ V}$  battery. It will draw a current of one ampere. When another resistor of  $2 \Omega$  is connected in parallel, it will also draw the same current. That is, total current drawn from the battery is  $2 \text{ A}$ . Hence, resistance of the circuit is halved. As we increase the number of resistors in parallel, the resistance of the circuit decreases and the current drawn from the battery goes on increasing.

In our homes, electrical appliances such as lamps, fans, heaters etc. are connected in parallel and each has a separate switch. Potential difference across each remains the same and their working is not influenced by others. As we switch on bulbs and fans, the resistance

of the electrical circuit of the house decreases and the current drawn from the mains goes on increasing (Fig. 19.10).



**Fig. 19.10 :** Arrangement of appliances in our homes. These are connected in parallel so that every appliance is connected to 220 V main supply. The total current drawn from the mains is the sum of the currents drawn by each appliance.

### Current division in parallel combination of two resistors

We know that  $I = I_1 + I_2$

but  $I_1 = \frac{V}{R_1}$  and  $I_2 = \frac{V}{R_2}$

substitute these values in above equation we have

$$I = \frac{V}{R_1} + \frac{V}{R_2}$$

$$I = V \left( \frac{R_1 + R_2}{R_1 R_2} \right)$$

or

$$V = I \frac{R_1 R_2}{R_1 + R_2}$$

Now current in  $R_1$  resistor

$$I_1 = \frac{V}{R_1} = I \frac{R_1 R_2}{R_1 + R_2} \times \frac{1}{R_1}$$

$$I_1 = \frac{(R_2)I}{R_1 + R_2} \quad (19.13 \text{ a})$$

Similarly

$$I_2 = \frac{V}{R_2} = I \frac{R_1 R_2}{R_1 + R_2} \times \frac{1}{R_2}$$

$$I_2 = \frac{(R_1)I}{R_1 + R_2} \quad (19.13 \text{ b})$$

In general, for current division in parallel combination of two resistors the current across any resistance equal to the opposite branch resistance divided by total resistance and multiplied by total current. **In general, for voltage division same resistance and for current division opposite branch resistance.**

### Example 19.5

For the circuit shown in Fig. 19.11, calculate the value of resistance  $R_2$ , and current  $I_2$  flowing through it.

#### Solution :

If the equivalent resistance of parallel combination of  $R_1$  and  $R_2$  is  $R$ , then

$$R = \frac{R_1 R_2}{R_1 + R_2} = \frac{10 R_2}{10 + R_2}$$

According to Ohm's law,

$$R = \frac{50}{10} = 5 \Omega$$

$$\therefore \frac{10 R_2}{10 + R_2} = 5$$

$$\Rightarrow 10 R_2 = 50 + 5 R_2 \quad \text{or} \quad R_2 = 10 \Omega$$

Since  $R_1$  and  $R_2$  are equal, current will be equally divided between them.

Hence,  $I_2 = 5A$

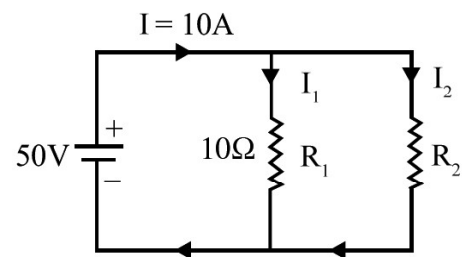


Fig. 19.11 : Two resistors in parallel

### Example 19.6

For the circuit shown in Fig. 19.12, calculate the equivalent resistance between points a and d.

#### Solution :

$15\Omega$  and  $3\Omega$  resistors are connected in parallel. The equivalent resistance of this combination is

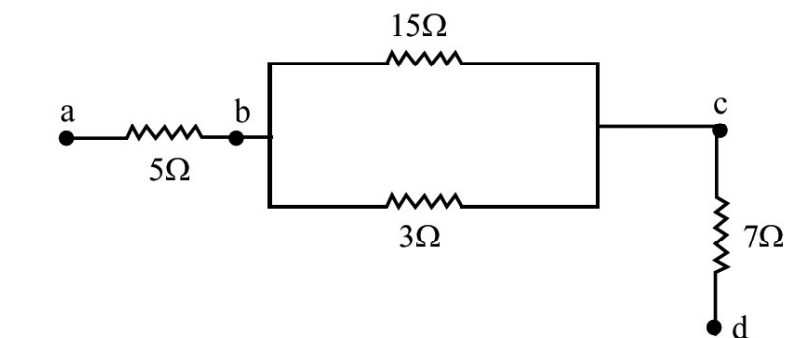


Fig. 19.12 : A combination of series and parallel groupings

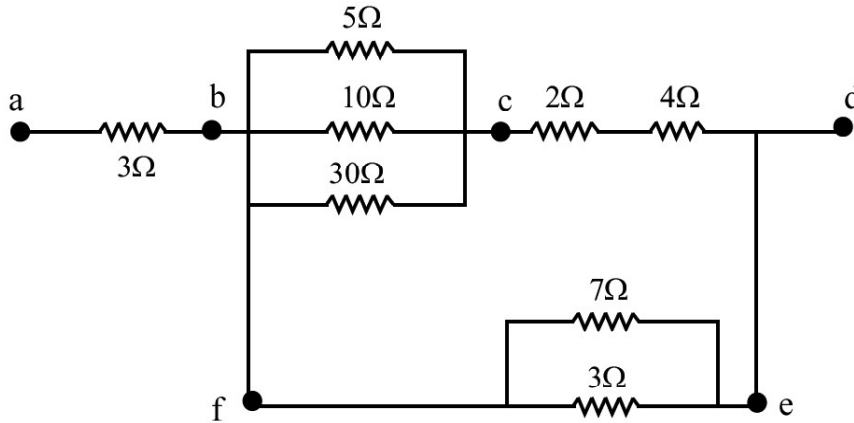
$$R_1 = \frac{15 \times 3}{15 + 3} \Omega = \frac{45}{18} = \frac{5}{2} = 2.5 \Omega$$

Now we can regard the resistances  $5\Omega$ ,  $R_1 = 2.5\Omega$  and  $7\Omega$  as connected in series. Hence, equivalent resistance between points a and d is

$$R = (5 + 2.5 + 7) = 14.5 \Omega$$

**Example 19.7**

Refer to the network shown in Fig. 19.13. Calculate the equivalent resistance between the points (i) b and c (ii) c and d, and (iii) a and e.



**Fig. 19.13** : A combination of series and parallel groupings of resistors

**Solution :**

- (i) Three resistors ( $5\Omega$ ,  $10\Omega$  and  $30\Omega$ ) are connected in parallel. Therefore, equivalent resistance is given by

$$\frac{1}{R_1} = \frac{1}{5} + \frac{1}{10} + \frac{1}{30} = \frac{6+3+1}{30} = \frac{10}{30}\Omega$$

or  $R_1 = 3\Omega$

- (ii) The resistors with resistances  $2\Omega$  and  $4\Omega$  are in series. The equivalent resistance

$$R_2 = (2 + 4) = 6\Omega$$

- (iii) The resistances  $7\Omega$  and  $3\Omega$  are in parallel. So equivalent resistance

$$\frac{1}{R_3} = \left( \frac{1}{7} + \frac{1}{3} \right) = \frac{3+7}{21} = \frac{10}{21}$$

or  $R_3 = \frac{21}{10}\Omega = 2.1\Omega$

Now we can treat equivalent resistance  $R_1$  and  $R_2$  to be in series. Therefore

$$R_4 = R_1 + R_2 = (3 + 6) = 9\Omega$$

Now  $R_4$  and  $R_3$  are in parallel. Therefore equivalent resistance

$$\begin{aligned} \frac{1}{R_5} &= \frac{1}{R_4} + \frac{1}{R_3} \\ &= \frac{1}{9} + \frac{1}{2.1} \end{aligned}$$

$$= \frac{1}{9} + \frac{10}{21} = \frac{37}{63}$$

$$R_5 = \frac{63}{57} \Omega = 1.70 \Omega$$

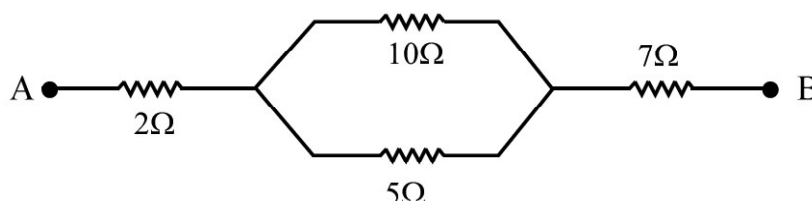
(iv) Finally  $R_5$  and  $3\Omega$  (between a and b) are in series. Hence

$$R = (1.70 + 3) = 4.70 \Omega$$

**Note :** For ease and convenience, you should draw a new equivalent circuit after every calculation.

### Intext Questions 19.2

1. There are two bulbs and a fan in your bed room. Are these connected in series or in parallel? Why?
2. The electric supply in a town is usually at 220 V. Sometimes the voltage shoots upto 300 V and may harm your TV set and other gadgets. What simple precaution can be taken to save your appliances?
3. Calculate the equivalent resistance between points A and B for the following circuit



4. In Function Halls several small bulbs are connected for decoration are these connected in parallel or series? Why?

### 19.4 TYPES OF RESISTORS

We use resistors in all electrical and electronic circuits to control the magnitude of current. Resistors usually are of two types :

- carbon resistors
- wire wound resistors

In a wire wound resistor, a resistance wire (of manganin, constantan or nichrome) of definite length, which depends on the required value of resistance, is wound two-fold over an insulating cylinder to make it non-inductive. In carbon resistors, carbon with a suitable binding agent is molded into a cylinder. Wire leads are attached to the cylinder for making connections to electrical circuits. Resistors are colour coded to give their values :

$$R = AB \times 10^C \Omega, D$$

where A, B and C are coloured stripes. The values of different colours are given in Table - 19.1. As may be noted,

- first two colours indicate the first two digits of the resistance value;
- third colour gives the power of ten for the multiplier of the value of the resistance; and
- fourth colour (D) (the last one) gives the tolerance of the resistance, which is 5% for golden colour, 10% for silver colour and 20% for body colour.

**Table - 19.1 : Colour codes of resistors**

Colour		Number	Multiplier
Black	B	0	1
Brown	B	1	$10^1$
Red	R	2	$10^2$
Orange	O	3	$10^3$
Yellow	Y	4	$10^4$
Green	G	5	$10^5$
Blue	B	6	$10^6$
Violet	V	7	$10^7$
Grey	G	8	$10^8$
White	W	9	$10^9$

For color code remember BBROY Goes to Bombay Via Gate Way.

Suppose that four colours on a resistor are Blue, Grey, Green and Silver. Then

The first digit will be 6 (blue)

The second digit will be 8 (Grey)

The third colour signifies multiplier  $10^5$  (Green)

The fourth colour defines tolerance = 10% (Silver)

Hence value of the resistance is

$$\begin{aligned}
 & 68 \times 10^5 \pm 10\% \\
 & = 68 \times 10^5 \pm (68 \times 10^5 \times 10/100) \\
 & = 68 \times 10^5 \pm 68 \times 10^4 \\
 & = (6.8 \pm 0.68) \text{ M}\Omega
 \end{aligned}$$

## 19.5) TEMPERATURE DEPENDENCE OF RESISTANCE

The resistivity of a conductor depends on temperature. For most metals, the resistivity increases with temperature and the change is linear over a limited range of temperature :

$$\rho = \rho_0 [1 + \alpha (T - T_0)] \quad (19.14)$$

where  $\rho$  and  $\rho_0$  are the resistivities at temperatures  $T$  and  $T_0$ , respectively. The temperatures are taken in  $^{\circ}\text{C}$  and  $T_0$  is the reference temperature.  $\alpha$  is called the temperature co-efficient of resistivity. Its unit is per degree celcius.

### Superconductors

Temperature dependence of resistivity led scientists to study the behaviour of materials at very low temperatures. They observed that certain metals and their alloys lost their resistivity completely below a certain temperature, called **transition temperature**, which is specific to the material. In such materials, current, once set up, remained, unchanged for ever without the use of an external source to maintain it. Such materials were termed as **superconductors**.

It was soon realised that superconductors, if they may exist near room temperature, will bring in revolutionary changes in technology. (These have been termed as *high temperature superconductors*.) For example, energy efficient powerful electromagnets made of superconducting coils may levitate vehicles above a magnetic track and make a high speed transportation system possible.

Efforts are being made to develop high temperature superconductors. The work done so far suggests that oxides of copper, barium and ytterium are showing good possibilities. A superconductor ( $\text{T}_2\text{Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_{10}$ ) which can exist at  $-153^{\circ}\text{C}$  has been developed. India is a front runner in this area of research.

Eqn. (19.14) can be rearranged to obtain an expression for temperature coefficient of resistivity :

$$\rho = \rho_0 + \rho_0 \alpha (T - T_0)$$

$$\alpha = \frac{(\rho - \rho_0)}{\rho_0 (T - T_0)} = \frac{1}{\rho_0} \frac{\Delta\rho}{\Delta T}$$

where  $\Delta\rho = (\rho - \rho_0)$  and  $\Delta T = T - T_0$

The resistivity versus temperature graph for a metal like copper is shown in Fig. 19.14 (a). The curve is linear over a wide range of temperatures.

You may recall that resistance of a conductor is proportional to its resistivity. Therefore, temperature variation of resistance can be written as :

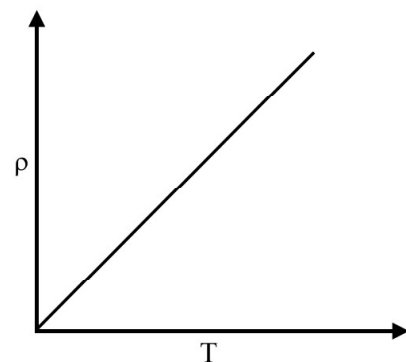


Fig. 19.14 : (a) Typical resistivity–temperature graph for a metal



$$R = R_0 [1 + \alpha (T - T_0)] \quad (19.15)$$

The resistances corresponding to two different temperatures  $T_1$  and  $T_2$  are given by

$$R_1 = R_0 [1 + \alpha (T_1 - T_0)] \quad (19.16)$$

and

$$R_2 = R_0 [1 + \alpha (T_2 - T_0)] \quad (19.17)$$

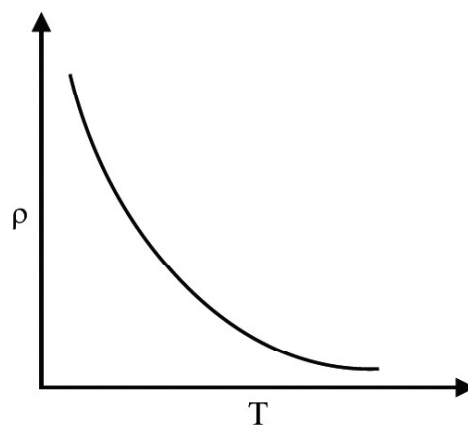
On combining these equations, we can write an expression for temperature coefficient of resistivity :

$$\alpha = \frac{(R_2 - R_1)}{R_0 (T_2 - T_1)} = \frac{1}{R_0} \frac{\Delta R}{\Delta T} \quad (19.18)$$

If  $R_0 = 1\Omega$  and  $(T_2 - T_1) = 1^\circ\text{C}$ , then  $\alpha = (R - R_0)$ . Thus **temperature coefficient of resistance is numerically equal to the change in resistance of a wire of resistance  $1\Omega$  at  $0^\circ\text{C}$  when the temperature changes by  $1^\circ\text{C}$ . This property of metals is used in making resistance thermometers named as RTD (Resistance temperature detectors).**

The resistivity of alloys also increases with increase in temperature. But the increase is very small compared to that for metals. For alloys such as **manganin, constantan and nichrome**, the temperature coefficient of resistivity is vanishingly small ( $\sim 10^{-6} \text{ }^\circ\text{C}^{-1}$ ) and resistivity is high. Hence these materials are used for making resistance wires or standard resistances.

Semiconductors such as germanium and silicon have resistivities which lie between those of metals and insulators. The resistivity of semiconductors usually decreases with increase in temperature [Fig.19.14 (b)]. This gives a negative temperature coefficient of resistance. This will be discussed in detail in the lesson on semiconductors.



**Fig. 19.14 (b) :** Resistivity of semiconductors decreases with temperature

## 19.6) ELECTROMOTIVE FORCE (emf) AND POTENTIAL DIFFERENCE

EMF is the short form of electromotive force. EMF of a cell or battery equals the potential difference between its terminals when these are not connected (open circuit) externally. You may easily understand the difference between e.m.f. and potential difference of a cell by performing the following activity.

### Activity 19.4

Connect a cell in a circuit having a resistor  $R$  and key  $K$ . A voltmeter of very high resistance is connected in parallel to the cell, as shown in Fig. 19.15 (a). When key  $K$  is closed, voltmeter reading will decrease. Can you give reasons for this decrease in the voltmeter

reading? Actually when key K is open, no current flows through the loop having cell and voltmeter: (The resistance in the circuit is infinite.) Hence the voltmeter reading gives e.m.f.  $E$  of the cell, which is the potential difference between the terminals of the cell when no current is drawn from it. When key K is closed, current flows outside and inside the cell. The cell introduces a resistance  $r$ , called **internal resistance** of the cell. Let current  $I$  be flowing in the circuit. Potential drop  $Ir$  across internal resistance  $r$  due to current flow acts opposite to the e.m.f. of the cell. Hence, the voltmeter reading will be

$$E - Ir = V$$

or

$$E = V + Ir \quad (19.19)$$

Thus while drawing current from a cell, e.m.f. of the cell is always greater than the potential difference across external resistance, unless internal resistance is zero. The internal resistance of an ideal cell is zero. When current enter into the cell as shown in Fig. 19.15 (b). The internal potential differences between the cell  $V_{ab} = E + Ir$  called as charging of the cell.

If the current is drawn from the cell as shown in Fig. 19.15 (c).

The terminal potential difference  $V_{ab} = E - Ir$  called as discharging of the cell.

E.M.F. of a cell depends on :

- the electrolyte used in the cell;
- the material of the electrodes; and
- the temperature of the cell.

Note that the e.m.f. of a cell does not depend on the size of the cell, i.e. on the area of plates and distance between them. This means that if you have two cells of different sizes, one big and one small, the e.m.f.s can be the same if the material of electrodes and electrolyte are the same. However, cells of larger size will offer higher resistance to the passage of current through it but can be used for a longer time.

### Example 19.8

When the current drawn from a battery is 0.5 A, potential difference at the terminals is 20 V. And when current drawn from it is 2.0 A, its voltage reduces to 16 V. Calculate the e.m.f. and internal resistance of the battery.

**Solution :**

Let  $E$  and  $r$  be the e.m.f. and internal resistance of battery. When current  $I$  is drawn from it, the potential drop across internal resistance of the cell is  $Ir$ . Then we can write

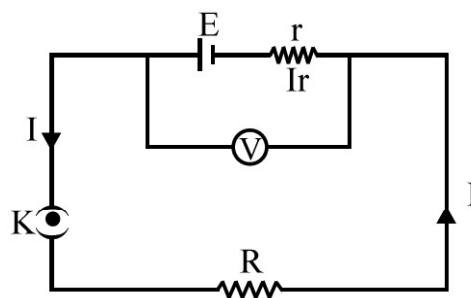


Fig. 19.15 (a)

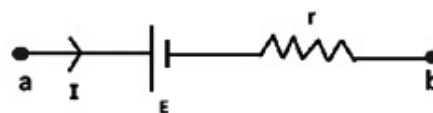


Fig. 19.15 (b)

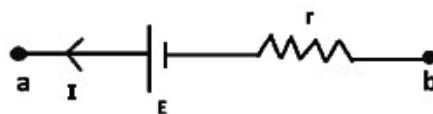


Fig. 19.15 (c)

$$V = E - Ir$$

For  $I = 0.5 \text{ A}$  and  $V = 20 \text{ volt}$ , we have

$$20 = E - 0.5 r \quad (\text{i})$$

For  $I = 2.0 \text{ A}$  and  $V = 16 \text{ volt}$ , we can write

$$16 = E - 2r \quad (\text{ii})$$

We can rewrite Eqns. (i) and (ii) as

$$2 E - r = 40$$

and

$$E - 2r = 16$$

Solving these, we get

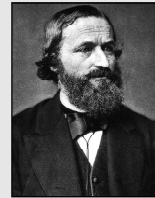
$$E = 21.3 \text{ V and } r = 2.67\Omega$$

## 19.7) KIRCHHOFF'S RULES

You now know that Ohm's law gives current–voltage relation for resistive circuits. But when the circuit is complicated, it is difficult to know current distribution by Ohm's law. In 1842, Kirchhoff formulated two rules which enable us to know the distribution of current in complicated electrical circuits or electrical networks.

### Gustav Robert Kirchhoff (1824-1887)

The fundamental contributions of German physicist Kirchhoff were in the fields of black body radiation and spectroscopy. But he also contributed in many other fields. His rules that you will study in this lesson enable us to analyse complex electric networks. With the help of Bunsen spectrum analysis, he discovered elements Rubidium and Cesium.



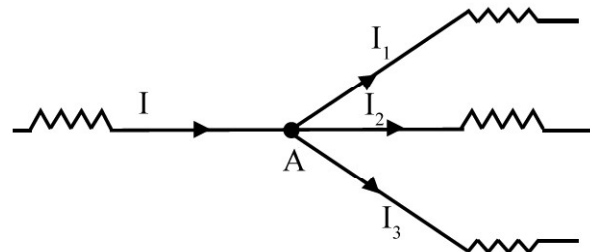
- (i) **Kirchhoff's First Rule (Junction Rule)** : It states that *the sum of all currents directed towards a junction (point) in an electrical network is equal to the sum of all the currents directed away from the junction*. It is based on the conservation of charge.

Refer to Fig. 19.16. If we take currents approaching point  $A$  as positive and those leaving it as negative, then we can write

$$I = I_1 + I_2 + I_3$$

$$\text{or } I - (I_1 + I_2 + I_3) = 0 \quad (19.20)$$

In other words, the algebraic sum of all currents at a junction is zero.



**Fig. 19.16** : Kirchhoff's first rule : Sum of currents coming to a junction is equal to the sum of currents going away from it.

Kirchhoff's first rule tells us that there is no accumulation of charge at any point if steady current flows in it. The net charge coming towards a point should be equal to that going away from it in the same time. In a way, it is an extension of continuity theorem in electrical circuits.

- (ii) **Kirchhoff's Second Rule (Loop Rule) :** This rule is an application of law of conservation of energy for electrical circuits. It tells us that *the algebraic sum of the products of the currents and resistances in any closed loop of an electrical network is equal to the algebraic sum of electromotive forces acting in the loop.* This is based on the conservation of energy.

While using this rule, we start from a point on the loop and go along the loop either clockwise or anticlockwise to reach the same point again. The product of current and resistance is taken as positive when we traverse in the direction of current. The e.m.f is taken positive when we traverse from negative to positive electrode through the cell. Mathematically, we can write

$$\Sigma IR = \Sigma E \quad (19.21)$$

Let us consider the electrical network shown in Fig. 19.17. For closed mesh ADCBA, we can write

$$I_1 R_1 - I_2 R_2 = E_1 - E_2$$

Similarly, for the mesh DHGCD

$$I_2 R_2 + (I_1 + I_2) R_3 = E_2$$

And for mesh AHGBA

$$I_1 R_1 + I_3 R_3 = E_1$$

At point D  $I_1 + I_2 = I_3$

In more general form, Kirchhoff's second rule is stated as : *The algebraic sum of all the potential differences along a closed loop in a circuit is zero.*

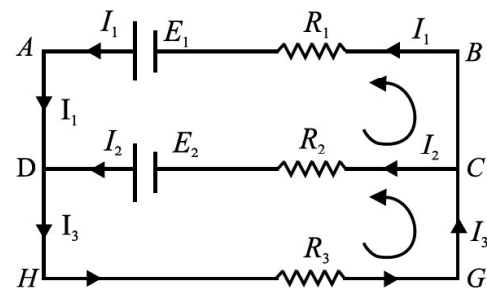


Fig. 19.17 : A network to illustrate Kirchhoff's second rule

### Example 19.9

Consider the network shown in Fig. 19.18. Current is supplied to the network by two batteries. Calculate the values of currents  $I_1$ ,  $I_2$  and  $I_3$ . The directions of the currents are as indicated by the arrows.

**Solution :**

Applying Kirchhoff's first rule to junction C, we get

$$I_1 + I_2 - I_3 = 0 \quad (i)$$

Applying Kirchhoff's second rule to the closed loops ACDBA and GCDFG, we get

$$5I_1 + 2I_3 = 12 \quad (ii)$$

and  $3I_2 + 2I_3 = 6 \quad (iii)$

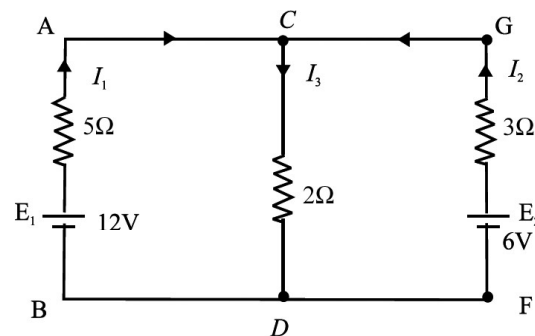


Fig. 19.18 : Calculation of currents in a network of resistors and batteries.

On combining these equations, we get

$$5I_1 - 3I_2 = 6 \quad (\text{iv})$$

Multiply (i) by 2 and add to (ii) to obtain

$$7I_1 + 2I_2 = 12 \quad (\text{v})$$

On multiplying Eqn. (iv) by 2 and Eqn. (v) by 3 and adding them, we get

$$31I_1 = 48$$

or

$$I_1 = 1.548 \text{ A}$$

Putting this value of  $I_1$  in eqn. (v), we get

$$I_2 = 0.582 \text{ A}$$

And from (i), we get

$$I_3 = I_1 + I_2 = 2.13 \text{ A}$$

### 19.7.1 Wheatstone Bridge

You have learnt that a resistance can be measured by Ohm's law using a voltmeter and an ammeter in an electrical circuit. But this measurement may not be accurate for low resistances. To overcome this difficulty, we use a wheatstone bridge. It is an arrangement of four resistances which can be used to measure one of them in terms of the other three.

Consider the circuit shown in Fig. 19.19 where

- (i) P and Q are two adjustable resistances connected in arms AB and BC.
- (ii) R is an adjustable known resistance.
- (iii) S is an unknown resistance to be measured.
- (iv) A sensitive galvanometer G along with a key  $K_2$  is connected in the arm BD.
- (v) A battery E along with a key  $K_1$  is connected in the arm AC.

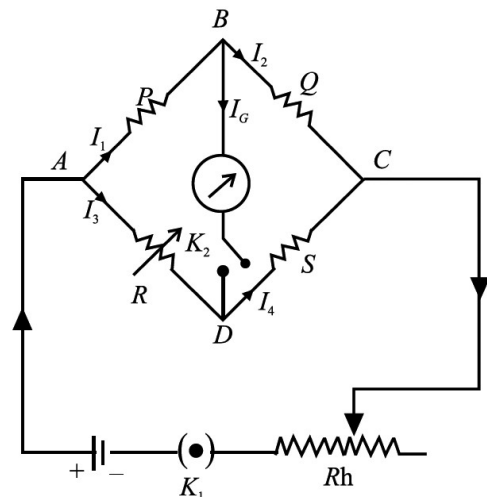


Fig. 19.19 : A wheatstone bridge

On closing the keys, in general, some current will flow through the galvanometer and you will see a deflection in the galvanometer. It indicates that there is some potential difference between points B and D. We now consider the following three possibilities:

- (i) **Point B is at a higher potential than point D :** Current will flow from B towards D and the galvanometer will show a deflection in one direction, say right
- (ii) **Point B is at a lower potential than point D :** Current will flow from point D towards B and the galvanometer will show a deflection in the opposite direction.
- (iii) **Both points B and D are at the same potential:** In this case, no current will flow through the galvanometer and it will show no deflection, i.e. the galvanometer is

in null condition. In this condition, the Wheatstone bridge is said to be in the *state of balance*.

The points B and D will be at the same potential only when the potential drop across P is equal to that across R. Thus

$$\text{But} \quad I_1 P = I_3 R \quad (19.22)$$

$$I_1 = I_2 + I_G$$

$$\text{and} \quad I_4 = I_3 + I_G \quad (19.23)$$

Applying Kirchhoff's first rule at junctions B and D in the null condition ( $I_G = 0$ ), we get

$$I_1 = I_2$$

$$\text{and} \quad I_3 = I_4 \quad (19.24)$$

Also potential drop across Q will be equal to that across S. Hence

$$I_2 Q = I_4 S \quad (19.25)$$

Dividing Eqn. (19.22) by Eqn. (19.25), we obtain

$$\frac{I_1 P}{I_2 Q} = \frac{I_3 R}{I_4 S} \quad (19.26)$$

Using Eqn. (19.24), we get

$$\frac{P}{Q} = \frac{R}{S} \quad (19.27)$$

This is the condition for which a Wheatstone bridge will be balanced. From Eqn. (19.27), we find that the unknown resistance S is given by

$$S = \frac{QR}{P}$$

In Fig. 19.19 the resistances are named as P, Q, R, S. In most textbooks follows same notation. If the positions of resistances P, Q, R, S are not properly placed as shown in Fig. 19.19. If they are placed differently, then the condition for bridge balance

$$\frac{P}{Q} = \frac{R}{S}$$

may not be the same. Hence it is always better to remember the product of opposite arm resistances are equal. Referring to Fig. 19.19

PS = RQ leads to

$$\frac{P}{Q} = \frac{R}{S}$$

Instead of resistances if four capacitances are connected in Wheatstone bridge, then bridge is balanced if the product of opposite arm reactances are equal.

You can easily see that measurement of resistance by Wheatstone bridge method has the following merits

- (i) *The balance condition given by Eqn. (19.27) at null position is independent of the applied voltage V. In other words, even if you change the e.m.f of the cell, the balance condition will not change.*
- (ii) *The measurement of resistance does not depend on the accuracy of calibration of the galvanometer. Galvanometer is used only as a null indicator (current detector).*

The main factor affecting the accuracy of measurement by Wheatstone bridge is its sensitivity with which the changes in the null condition can be detected. It has been found that the bridge has the greatest sensitivity when the resistances in all the arms are nearly equal.

**Example 19.10**

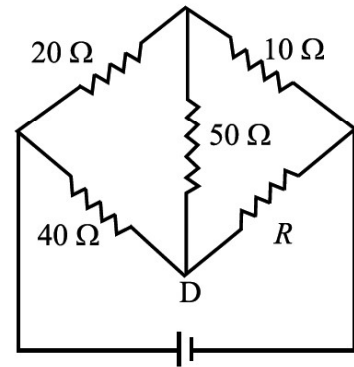
Calculate the value of R shown in Fig. 19.20 when there is no current in 50Ω resistor.

**Solution :**

This is Wheatstone bridge where galvanometer has been replaced by 50Ω resistor. The bridge is balanced because there is no current in 50Ω resistor. Hence,

$$\frac{20}{10} = \frac{40}{R}$$

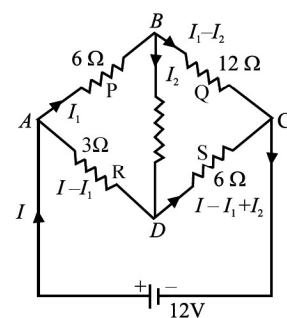
or 
$$R = \frac{40 \times 10}{20} = 20\Omega$$



**Fig. 19.20 :** When there is no current through 50Ω resistor, the bridge is balanced.

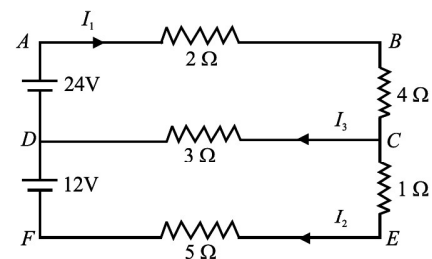
**Intext Questions 19.3**

1. Refere to figure below. Calculate the value of currents in the arms AB, AD and BD.



2. Examine the following circuit resistors and batteries. Calculate the current.

$I_1, I_2$  and  $I_3$ .



## 19.8) POTENTIOMETER

You now know how to measure e.m.f. of a source or potential difference across a circuit element using a voltmeter. (An ideal voltmeter should have infinite resistance so that it does not draw any current when connected across a source of e.m.f.) Practically it is not possible to manufacture a voltmeter which will not draw any current. To overcome this difficulty, we use a potentiometer, which draws no current from it. It employs a null method. The potentiometer can also be used for the measurement of internal resistance of a cell, the current flowing in a circuit and comparison of resistances.

### 19.8.1 Description of a Potentiometer

A potentiometer consists of a wooden board on which a number of resistance wires (usually ten) of uniform cross-sectional area are stretched parallel to each other. The wire is of manganin or nichrome. These wires are joined in series by thick copper strips. In this way, these wires together act as a single wire of length equal to the sum of the lengths of all the wires. The end terminals of the wires are provided with connecting screws.

A metre scale is fixed on the wooden board parallel to wires. A jockey (a sliding contact maker) is provided with the arrangement. It makes a knife edge contact at any desired point on a wire. Jockey has a

pointer which moves over the scale. It determines the position of the knife edge contact. In Fig. 19.21 a ten wire potentiometer is shown. A and B are ends of the wire. K is a jockey and S is a scale. Jockey slides over a rod CD.

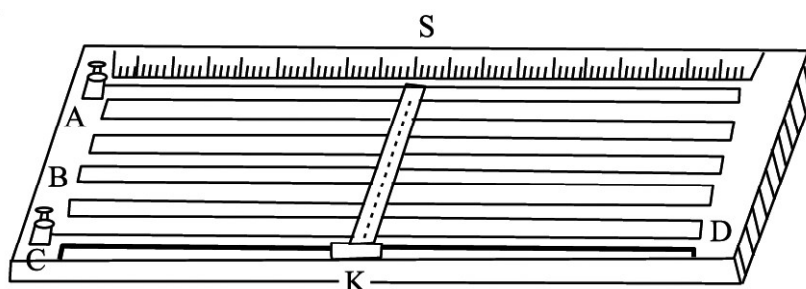


Fig. 19.21 : An illustrative diagram of a potentiometer

### 19.8.2 Measurements with a Potentiometer

Let us suppose that a steady source of e.m.f.  $E$  (say an accumulator) is connected across a uniform wire  $AB$  of length  $l$ . Positive terminal of the accumulator is connected at end  $A$  (Fig. 19.22). A steady current  $I$  flows through the wire. The potential difference across  $AB$  is given by

$$V_{AB} = RI$$

If  $r$  is the resistance per unit length of the wire, and  $k$  is the potential drop across unit length of the wire, then

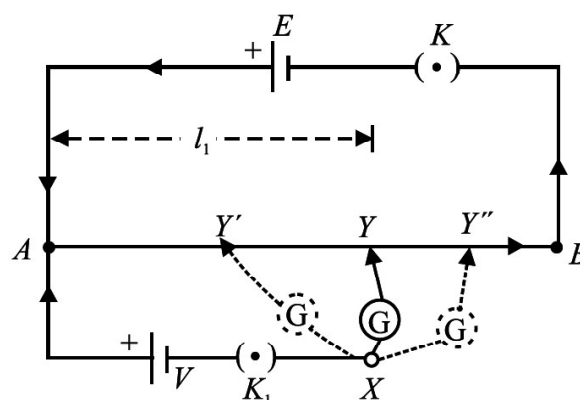


Fig. 19.22 : Potentiometer circuit to measure potential difference between the terminals of a cell.



$$R = r/l$$

and 
$$E = kl$$

or 
$$k = \frac{E}{l}$$

For length  $l_1$  of wire, potential drop is given by

$$V_1 = kl_1 = \frac{E}{l}l_1 \quad (19.28)$$

Thus potential falls linearly with distance along the wire from the positive to the negative end.

We wish to measure an unknown voltage  $V$ . The positive terminal of the cell is connected to end A of the wire and negative terminal through a galvanometer to the jockey having variable contact Y. Note that for  $V > E$ , it will not be possible to obtain a null point. So we use a standard cell of emf  $E (> V)$ , as shown in Fig. 19.22. To check this, insert keys  $K$  and  $K_1$  and tap at ends A and B. The galvanometer should show deflection in opposite directions. If so, all is well with the circuit.

Insert key  $K_1$  and start moving jockey from A towards B. Suppose that at position  $Y'$  potential drop across the length  $AY'$  of the wire is less than voltage  $V$ . The current in the loop  $AY'XA$  due to voltage  $V$  exceeds the current due to potential difference across  $AY'$ . Hence galvanometer shows some deflection in one direction. Then jockey is moved away, say to  $Y''$  such that potential drop across  $AY''$  is greater than the voltage  $V$ . If galvanometer shows deflection in the other direction, the voltage drop across  $AY''$  is greater than that across  $AY'$ . Therefore, the jockey is moved slowly between  $Y'$  and  $Y''$ . A stage is reached, say at point Y, where potential drop across  $AY$  is equal to voltage  $V$ . Then points X and Y will be at the same potential and hence the galvanometer will not show any deflection, i.e. null point is achieved. If  $l_1$  is the length between A and Y, then

$$V = kl_1 = \frac{El_1}{l} \quad (19.29)$$

Thus, the unknown voltage  $V$  is measured when no current is drawn.

The measurements with potentiometer have following advantages :

- When the potentiometer is balanced, no current is drawn from the circuit on which the measurement is being made.
- It produces no change in conditions in a circuit to which it is connected.
- It makes use of null method for the measurement and the galvanometer used need not be calibrated.

### 19.8.3 Comparison of E.M.Fs of two Cells

You have learnt to measure the e.m.f. of a cell using a potentiometer. We shall now extend the same technique for comparison of e.m.fs of two cells. Let us take, for example, a Daniel cell and a Leclanche cell and let  $E_1$  and  $E_2$  be their respective e.m.fs.

Refer to circuit diagram shown in Fig. 19.23. The cell of e.m.f.  $E_1$  is connected in the circuit through terminals 1 and 3 of key  $K_1$ . The balance point is obtained by moving the jockey on the potentiometer wire as explained earlier. Note that e.m.f of cell  $E$  should be greater than the emfs of  $E_1$  and  $E_2$  separately. (Otherwise, balance point will not be obtained.) Let the balance point on potentiometer be at point  $Y_1$  and length  $AY_1 = l_1$ . The cell of e.m.f.  $E_2$  is connected in the circuit through terminals 2 and 3 of the key  $K_2$ . Suppose balance is obtained at point  $Y_2$  and length  $AY_2 = l_2$ .

Applying potentiometer principle, we can write

$$E_1 = kl_1 \quad \text{and} \quad E_2 = kl_2$$

where  $k$  is the potential gradient along the wire  $AB$ . Hence

$$\frac{E_1}{E_2} = \frac{l_1}{l_2} \quad (19.30)$$

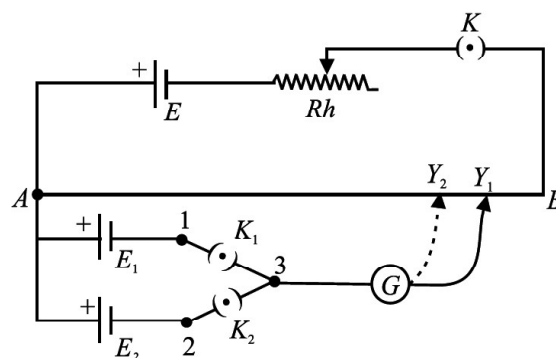


Fig. 19.23 : Circuit diagram for comparison of e.m.f.s of two cells  $E_1$  and  $E_2$ .

#### 19.8.4 Determination of Internal Resistance of a Cell

You have learnt that cells always offer resistance to the flow of current through them, which is often very small. This resistance is called the internal resistance of the cell and depends on the size of the cell, i.e. the area of the plates immersed in the liquid, the distance between the plates and strength of electrolyte used in the cell.

Let us now learn how to measure internal resistance of a cell using a potentiometer. Refer to Fig. 19.24, which shows the circuit diagram for measuring internal resistance ' $r$ ' of a cell of emf  $E_1$ . A resistance box  $R$  with a key  $K_1$  is connected in parallel with the cell. The primary circuit has a standard cell, a rheostat and a one way key  $K$ . As soon as key  $K$  is closed, a current  $I$  begins to flow through the wire  $AB$ . The key  $K_1$  is kept open and on moving the jockey, a balance is obtained with the cell  $E_1$  at point, say  $Y_1$ . Let  $AY_1 = l_1$ . Then we can write

$$E_1 = kl_1 \quad (19.31 \text{ a})$$

Now key  $K_1$  is closed. This introduces a resistance across the cell. A current, say  $I_1$ , flows in the loop  $E_1RK_1E_1$  due to cell  $E$ . Using Ohm's law, we can write

$$I_1 = \frac{E_1}{R+r}$$

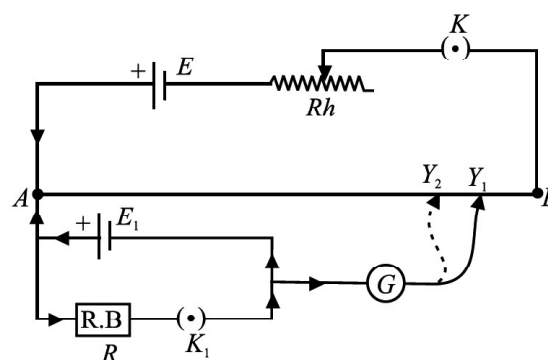


Fig. 19.24 : Measurement of the internal resistance  $r$  of a cell

where  $r$  is internal resistance of the cell. It means that terminal potential difference  $V_1$  of the cell will be less than  $E_1$  by an amount  $I_1 r$ . The value of  $V_1$  is

$$V_1 = I_1 R = \frac{E_1}{R+r} R$$

Then, potential difference  $V_1$  is balanced on the potentiometer wire without change in current  $I$ . Let the balance point be at point  $Y_2$  such that  $AY_2 = l_2$ . Then

$$V_1 = k l_2 \quad (19.31 \text{ b})$$

Using Eqns. (19.31 a, b) we get

$$\frac{E_1}{V_1} = \frac{l_1}{l_2} = \frac{R+r}{R}$$

or 
$$r = R \left( \frac{l_1}{l_2} - 1 \right) \quad (19.32)$$

Thus by knowing  $R_1$ ,  $l_1$  and  $l_2$ , the value of  $r$  can be easily calculated.

### Example 19.11

Length of a potentiometer wire is 5 m. It is connected with a battery of fixed e.m.f. Null point is obtained for the Daniel cell at 100 cm. If the length of the wire is kept 7 m, what will be the position of null point?

#### Solution :

Let e.m.f. of battery be  $E$  volt. The potential gradient for 5 m length is

$$k_1 = \frac{E}{5} \text{Vm}^{-1}$$

When the length of potentiometer wire is 7 m, potential gradient is

$$k_2 = \frac{E}{7} \text{Vm}^{-1}$$

Now, if null point is obtained at length  $l_2$ , then

$$E_1 = k_2 l_2 = \frac{E}{7} l_2$$

Here same cell is used in two arrangements. Hence

$$\frac{E}{5} = \frac{E}{7} l_2$$

$$\Rightarrow l_2 = 7 / 5 = 1.4 \text{ m}$$

## 19.9) DRIFT VELOCITY OF ELECTRONS

Let us now understand the microscopic picture of electrical conduction in a metal. The model presented here is simple but its strength lies in the fact that it conforms to Ohm's law.

We assume that a metallic solid consists of atoms arranged in a regular fashion. Each atom usually contributes free electrons, also called conduction electrons. These electrons are free to move in the metal in a random manner, almost the same way as atoms or molecules of a gas move about freely in a container. It is for this reason that sometimes conduction electrons are referred to as **electron gas**. The average speed of conduction electrons is about  $10^6 \text{ ms}^{-1}$ .

We know that no current flows through a conductor in the absence of an electric field, because the **average velocity** of free electrons is zero. On an average, the number of electrons moving in  $+x$  direction is same as number of electrons moving in  $-x$  direction. Thus there is no net flow of charge in any direction.

The conduction electrons frequently collide with the atoms in the solid. The free electrons drift slowly in a direction opposite to the direction of the applied electric field. The average drift velocity is of the order of  $10^{-4} \text{ ms}^{-1}$ . This is very small compared to the average speed of free electrons between two successive collisions ( $10^6 \text{ ms}^{-1}$ ). On applying an electric field, the conduction electrons get accelerated. The excess energy gained by the electrons is lost during collisions with the atoms. The atoms gain energy and vibrate more vigorously. The conductor gets heated up. Fig. 19.25 shows how the motion of electrons is modified when an electric field is applied.

Let us now obtain an expression for the drift velocity of conduction electrons. Let  $e$  and  $m$  be the charge and mass of an electron respectively. If  $E$  is the electric field, the force on the electron is  $eE$ . Hence acceleration experienced by the electron is given by

$$\mathbf{a} = \frac{e\mathbf{E}}{m}$$

If  $\tau$  is the average time between collisions, we can write the expression for drift velocity of an electrons in terms of electric field as

$$v_d = \frac{e\mathbf{E}}{m} \tau$$

On combining this result with Eqn. (19.4), we obtain the expression for current :

$$I = -neAv_d$$

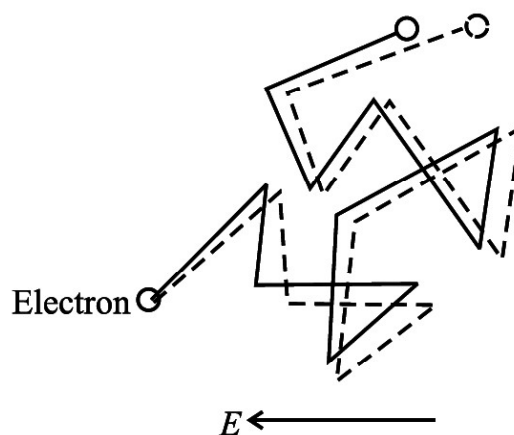


Fig. 19.25 : Motion of electrons in a conductor placed in an electric field.

$$= -neA \frac{eE}{m} \tau$$

$$= -\frac{Ane^2E}{m} \tau$$

Since electric field is negative spatial gradient of potential  $\left(E = -\frac{\partial V}{\partial r}\right)$  we can rewrite the expression for current as

$$I = +\frac{ne^2A}{m} \frac{V}{l} \tau \quad (19.33)$$

$$\Rightarrow \frac{V}{I} = \frac{m}{ne^2\tau} \frac{l}{A} = R \quad (19.34)$$

Eqn. (19.34) implies that conduction current obeys Ohm's law.

On combining this result with Eqn. (19.9), we get

$$\rho = \frac{1}{\sigma} = \frac{m}{ne^2\tau} \quad (19.35)$$

## Electrical Energy

We know that when current is passed through a resistor heat developed in the resistor is actually due to the work done by the seat of e.m.f (the electric cell). The electric cell is capable of doing the work from its stored chemical energy. Thus, an electric cell converts chemical energy into electrical energy.

Let us suppose that as the electric cell drives a charge  $q$  through a resistor of resistance  $R$  in a time  $t$ , the current passing through the resistor is  $i$ . As a result, a potential difference  $V$  will be developed across the resistor. The electrical work ( $W$ ) done by the cell in driving the charge  $q$  through a p.d.  $V$  against the resistance of the resistor will be

$$W = qV$$

The charge  $q$  and current  $i$  are related to each other by  $i = \frac{q}{t}$  or  $q = it$ .

Hence, we have  $W = V it$ .

Again, from Ohm's law we have  $V = iR$ .

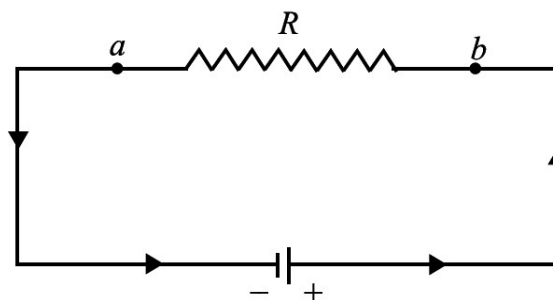
$$W = Vit = i^2Rt = \frac{V^2}{R} t$$

The unit of electrical work is joule.

$$\text{Electrical energy } W = Vit = i^2Rt = \frac{V^2}{R}t$$

### 19.10 POWER CONSUMED IN AN ELECTRICAL CIRCUIT

Let us examine the circuit in Fig. 19.26 where a battery is connected to an external resistor  $R$ . The positive charges (so to say) flow in the direction of the current in the resistor and from negative to positive terminal inside the battery. The potential difference between two points gives kinetic energy to the charges. These moving charges collide with the atoms (ions) in the resistor and thus lose a part of their kinetic energy. This energy increases with the temperature of the resistor. The loss of energy by moving charges is made up at the expense of chemical energy of the battery.



**Fig. 19.26 :** A circuit containing a battery and a resistor. The power consumed depends on the potential difference between the points  $a$  and  $b$ , the current through the resistor.

The rate of loss of potential energy by moving charge  $\Delta Q$  in going through the resistor is

$$\frac{\Delta U}{\Delta t} = V \frac{\Delta Q}{\Delta t} = VI \quad (19.36)$$

where  $I$  is the current in the circuit and  $V$  is potential difference between the ends of the resistor.

It is assumed that the resistance of the connecting wires is negligible. The total loss is in the resistor  $R$  only. Rate of loss of energy is defined as power :

$$P = VI$$

Since  $V = IR$ , we can write

$$P = I^2R = V^2/R \quad (19.37)$$

The SI unit of power is watt (W).

The electrical power lost in a conductor as heat is called *joule heat*. The heat produced is proportional to : (i) square of current ( $I$ ), (ii) resistance of conductor ( $R$ ), and (iii) time for which current is passed ( $t$ ).

The statement  $Q = I^2 Rt$ , is called Joule's law for heating effect of current.

#### Example 19.12

A 60W lamp is connected to 220 V electricity supply in your home. Calculate the power consumed by it, the resistance of its filament and the current through it.

**Solution :**

We know that  $I = P/V$

$$\therefore = \frac{60 \text{ W}}{220 \text{ V}} = \frac{3}{11} = 0.27 \text{ A}$$

Resistance of the lamp

$$\begin{aligned} R &= \frac{V}{I} \\ &= \frac{220 \text{ V}}{3/11 \text{ A}} \\ &= \frac{220 \times 11}{3} \Omega = 807 \Omega \end{aligned}$$

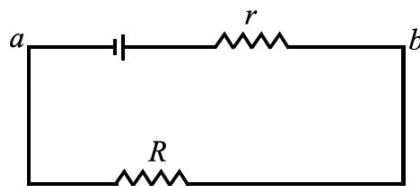
The lamp consumes 60 J of energy per second. It will consume 60 Wh energy in one hour and  $60 \times 24 = 1440$  Wh energy in one day.

Energy consumed per day = 1.440 kWh

In common man's language, it is known as 1.4 unit of energy.

**Intext Questions 19.4**

- When current drawn from a cell increases, the potential difference between the cell electrodes decreases. Why?
- A metallic wire has a resistance of  $30 \Omega$  at  $20^\circ \text{C}$  and  $30.16 \Omega$  at  $40^\circ \text{C}$ . Calculate the temperature coefficient of resistance.
- The e.m.f of a cell is 5.0 V and R in the circuit is  $4.5 \Omega$ . If the potential difference between the points a and b is 3.0 V, calculate the internal resistance r of the cell.



- In a potentiometer circuit, balance point is obtained at 45 cm from end A when an unknown e.m.f is measured. The balance point shifts to 30 cm from this end when a cell of 1.02 V is put in the circuit. Standard cell E always supplies a constant current. Calculate the value of unknown e.m.f.
- A potentiometer circuit is used to compare the e.m.f. of two cells  $E_1$  and  $E_2$ . The balance point is obtained at lengths 30 cm and 45 cm, respectively for  $E_1$  and  $E_2$ . What is the e.m.f of  $E_1$ , if  $E_2$  is 3.0 V?
- A current of 0.30 A flows through a resistance of  $500 \Omega$ . How much power is lost in the resistor?

7. You have two electric lamps. The printed specifications on them are 40 W, 220 V and 100 W, 220 V. Calculate the current and resistance of each lamp when put in a circuit of 220 V supply line.

## WHAT YOU HAVE LEARNT

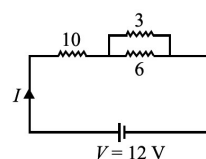
- Drift velocity is the average velocity with which an electrons move opposite to the field when an electric field exists in a conductor.
- Electric current through any cross-sectional area is the rate of transfer of charge from one side to other side of the area. Unit of current is ampere and is denoted by A.
- Ohm's law states that the current flowing through a conductor is proportional to the potential difference when physical conditions like pressure and temperature remain unchanged.
- Ratio  $V/I$  is called resistance and is denoted by R. Unit of resistance is ohm (denoted by  $\Omega$ )
- Resistivity (or specific resistance) of a material equals the resistance of a wire of the material of one metre length and one  $m^2$  area of cross section. Unit of resistivity is ohm metre.
- For a series combination of resistors, the equivalent resistance is sum of resistances of all resistors.
- For a parallel combination of resistors, inverse of equivalent resistance is equal to the sum of inverses of all the resistances.
- Kirchhoff's rules help us to study systematically the complicated electrical circuits. The first rule states that the sum of all the currents directed towards a point in an electrical network is equal to the sum of all currents directed away from the point. Rule II : The algebraic sum of all potential differences along a closed loop in an electrical network is zero.
- The Wheatstone bridge circuit is used to measure accurately an unknown resistance (S) by comparing it with known resistance (P, Q and R). In the balance condition,  $P/Q = R/S$ .
- The e.m.f. of a cell is equal to the potential difference between its terminals when a circuit is not connected to it.
- A potentiometer measures voltages without drawing current. Therefore, it can be used to measure e.m.f. of a source that has appreciable internal resistance.
- Drift velocity of electrons in a conductor is given by  $v_d = -\frac{eE}{m} \tau$ .
- Power consumed in an electrical circuit through Joule heating is given by

$$p = VI = I^2R = \frac{V^2}{R}.$$

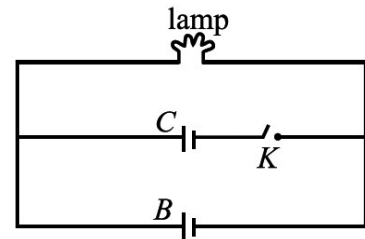


### TERMINAL EXERCISE

1. Explain the drift of free electrons in a metallic conductor under external electric field. Derive an expression for drift velocity.
  2. Define electric current and discuss Ohm's law.
  3. Define resistivity of a conductor. How does the resistance of a wire depend on the resistivity of its material, its length and area of cross-section?
  4. Define electrical conductivity. Write its unit. How does electrical conductivity depend on free electron concentration of the conductor?
  5. Explain the difference between ohmic and non-ohmic resistances. Give some examples of non-ohmic resistances.
  6. What is the effect of temperature on the resistivity of a material? Why does electrical conductivity of a metal decrease with increase in temperature?
- 
7. The colours on the resistor shown here are red, orange, green and gold as read from left to right. How much is the resistance according to colour code?
  8. Three resistors of resistances  $R_1$ ,  $R_2$  and  $R_3$  are connected (i) in series, and (ii) in parallel. Calculate the equivalent resistance of combination in each case.
  9. What is the difference between emf and potential difference between the electrodes of a cell. Derive relation between the two.
  10. State Kirchhoff's rules governing the currents and electromotive forces in an electrical network?
  11. Give theory of Wheatstone's bridge method for measuring resistances.
  12. Discuss the theory of potentiometer.
  13. How will you measure unknown potential difference with the help of a potentiometer?
  14. Describe potentiometer method of comparing e.m.f. of two cells.
  15. How will you determine internal resistance of a cell with the help of a potentiometer? What factors are responsible for internal resistance of a cell?
  16. A wire of length 1 m and radius 0.1 mm has a resistance of  $100\Omega$ . Calculate the resistivity of the material.
  17. Consider a wire of length 4 m and cross-sectional area  $1\text{ mm}^2$  carrying a current of 2 A. If each cubic meter of the material contains  $10^{29}$  free electrons, calculate the average time taken by an electron to cross the length of the wire.
  18. Suppose you have three resistors, each of value  $30\Omega$ . List all the different resistances that you can obtain by combining them.
  19. The potential difference between the terminals of a battery of e.m.f. 6.0 V and internal resistance  $1\Omega$  drops to 5.8 V when connected across an external resistor. Find the resistance of the external resistor.



20. For the circuit shown here, calculate the value of current  $I$  and equivalent resistance  $R$ .

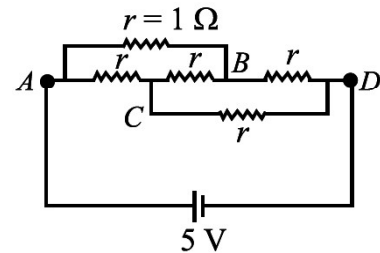


21. Examine the following network containing a lamp, a capacitor and a battery. The lamp is lighted when connected directly to the battery. What happens to it in this circuit when the switch is closed.

22. The following Wheatstone's bridge is balanced.

Calculate

- (a) the value of equivalent resistance  $R$  in the circuit, and  
 (b) the current in the arms  $AB$  and  $DC$ .



## ANSWERS TO INTEXT QUESTIONS

### 19.1

- (a) The current reduces to half as resistance of the wire is doubled.  
 (b) The current is doubled as resistance is halved.
- Resistivity is a property of the material of wire. It will not change with change in length and area of cross-section.

$$\rho = 2 \times 10^{-8} \Omega \text{m}$$

$$3. \quad R = \frac{V}{I} = \frac{8}{0.15} = \frac{800}{15} = 53.3 \Omega$$

$$R = \frac{\rho l}{A} \Rightarrow \frac{800}{15} = \rho \frac{3}{2 \times 10^{-4}} \Rightarrow \rho = \frac{800 \times 2 \times 10^{-4}}{15 \times 3} = 35.5 \times 10^{-4} \Omega \text{m}$$

- No. Only metallic conductor obey Ohm's law upto a certain limit. Semiconductors and electrolytes do not obey Ohm's law.

$$5. \quad I = \frac{q}{\tau} = \frac{n|e|}{t} = \frac{5 \times 10^{17} \times 1.6 \times 10^{-19}}{1} \text{ A} = 0.8 \times 10^{-3} \text{ A} = 0.8 \text{ mA}$$

The direction of current is opposite to the direction of flow of electrons, i.e., from right to left.

## 19.2

1. In parallel. They may draw different currents needed for their operation and we operated separately using different switches.
2. We use a voltage stabilizer.
3. 
$$R = R_1 + \frac{R_2 R_3}{R_2 + R_3} + R_4$$
$$= 2 + \frac{10}{3} + 7$$
$$= 12.3 \Omega$$
4. In function halls small bulbs are connected in series.

## 19.3

$$1. \quad \frac{P}{Q} = \frac{6}{12} = \frac{1}{2} \quad \text{and} \quad \frac{R}{S} = \frac{3}{6} = \frac{1}{2}$$

$$\frac{P}{Q} = \frac{R}{S} \quad \therefore \text{bridge is balanced}$$

$$\text{Hence } V_B = V_D \text{ and } I_2 = 0$$

$$I_1 = \frac{V}{R} = \frac{12}{18} = \frac{2}{3} \text{ A}$$

and

$$I - I_1 = \frac{12}{9} = \frac{4}{3} \text{ A}$$

2. Applying Kirchoff's second rule on loop ABCDA, we get

$$2I_1 + 4I_1 + 3I_3 = 24$$

$$6I_1 + 3I_3 = 24 \quad \dots (1) \quad \Rightarrow \quad 2I_1 + I_3 = 8 \quad \dots (1)$$

Similarly, for loop DCEFD, we can write

$$-3I_3 + 6I_2 = 12 \quad \Rightarrow \quad 2I_2 - I_3 = 4 \quad \dots (2)$$

Also applying Kirchoff's first rule at junction D we get

$$I_2 + I_3 = I_1$$

Substituting in (1) we get

$$2I_2 + 3I_3 = 8$$

$$2I_2 - I_3 = 4$$

$$4I_3 = 4$$

$$I_3 = 1\text{A}$$

Substituting in (2)

$$2I_2 = 5 \quad \Rightarrow \quad I_2 = 2.5\text{ A}$$

## 19.4

1.  $V = E - Ir$  as  $I$  increases  $V$  decreases.

$$2. \quad R_{20} = R_0 (1 + 20 \alpha)$$

$$R_{40} = R_0 (1 + 40 \alpha)$$

$$\frac{R_{40}}{R_{20}} = \frac{1 + 40\alpha}{1 + 20\alpha}$$

$$\frac{1 + 40\alpha}{1 + 20\alpha} = \frac{30.16}{30} = 1 + \frac{0.16}{30}$$

$$1 + \frac{20\alpha}{1 + 20\alpha} = 1 + \frac{0.16}{30}$$

On cross-multiplication, we get  $600 \alpha = 0.16 + 3.2 \alpha$

$$\Rightarrow \alpha \approx \frac{0.16}{600} = 2.67 \times 10^{-4} \text{ K}^{-1}$$

$$3. \quad I = \frac{V}{R} = \frac{3}{4.5} = \frac{30}{45} = \frac{2}{3} \text{ A}$$

$$V = E - Ir \Rightarrow 3 = 5 - \frac{2}{3}r$$

$$\therefore r = \frac{2 \times 3}{2} = 3\Omega$$

$$4. \quad \frac{E_2}{E_1} = \frac{l_2}{l_1} \Rightarrow \frac{1.02}{E_1} = \frac{30}{45} \Rightarrow E_1 = 0.51 \times \frac{3}{2} = 1.53 \text{ V}$$

$$5. \quad \frac{E_2}{E_1} = \frac{l_2}{l_1}$$

$$\frac{E_1}{3} = \frac{2}{3}$$

$$E_1 = 2 \text{ V}$$

$$\begin{aligned} 6. \quad P &= I^2 R \\ &= 0.3 \times 0.3 \times 500 \\ &= 45 \text{ W.} \end{aligned}$$

$$7. \quad I = \frac{P}{V} \Rightarrow I_1 = \frac{40}{220} = \frac{2}{11} \text{ A} \quad \text{and} \quad I_2 = \frac{100}{220} = \frac{5}{11} \text{ A}$$

$$R = \frac{V^2}{P}, \Rightarrow R_1 = \frac{220 \times 220}{40} = 1210 \Omega$$

$$\text{and} \quad R_2 = \frac{220 \times 220}{100} = 484 \Omega$$

### ANSWERS TO TERMINAL EXERCISE

16.  $3.14 \times 10^{-6} \Omega\text{m}$

17. 32 ms.

18. (i) All resistance in series; equivalent resistance  $90 \Omega$

(ii) All resistance in parallel; equivalent resistance  $10 \Omega$

(iii) One resistance in series with two others which are connected in parallel; equivalent resistance  $45 \Omega$

(iv) Two resistance in series and one resistance in parallel to them; equivalent resistance  $20 \Omega$

19.  $29 \Omega$

20.  $I = 1 \text{ A}, R = 12 \Omega$

22. (a)  $R = r = 1 \Omega$

(b)  $I = 2.5 \text{ A}$



# MAGNETISM

## INTRODUCTION

In the previous chapter, electric charges and fields, we learnt how charged rods attract each other or small bits of paper. You might have also played with magnets – the substances having the property of attracting small bits of iron. The phenomenon of magnetism was known to Greeks as early as 600 B.C. They observed that some stones called magnetite or lodestone ( $\text{Fe}_3\text{O}_4$ ) attracted iron pieces. The pieces of naturally occurring magnetite are called natural magnets. The word magnet is derived from the name of an island in Greece called magnesia where magnetic ore deposits were found.

### OBJECTIVES

After studying this lesson, you should be able to

- *know what are the magnets and their properties;*
- *define magnetic field and field lines;*
- *describe the effect of uniform magnetic field on a suspended magnet (or dipole);*
- *define the magnetic flux, magnetic field induction and their units;*
- *understand the Gauss's law in magnetism;*
- *define the magnetisation and magnetic intensity;*
- *explain the magnetic properties of different magnetic materials;*
- *list the elements of earth's magnetic field.*

### 20.1 MAGNETS AND THEIR PROPERTIES

The property of any object by virtue of which it can attract a piece of iron (or steel or cobalt or nickel) is called magnetism. A natural magnet is an ore of iron ( $\text{Fe}_3\text{O}_4$ ), which attracts small pieces of iron, cobalt and nickel towards it. Natural magnets are weak, but materials like iron, nickel, cobalt may be converted into strong permanent magnets which are known as artificial magnets, e.g., a bar magnet, a magnetic needle, a horse-shoe magnet etc. All magnets– natural or artificial, have same properties. You must be familiar with basic properties of magnets. However, for completeness, we recapitulate these.

1. **Directive Property** : A small bar magnet, when suspended freely on its center of mass so as to rotate about a vertical axis, always stays in approximately geographical north-south direction.

2. **Attractive Property :** A magnet attracts small pieces of magnetic materials like iron, nickel and cobalt. The force of attraction is maximum at points near the ends of the magnet. These points are called poles of the magnet. In a freely suspended magnet, the pole which points towards the geographical north is called north pole and the one which points towards the geographical south is called south pole. Do directive and attractive properties suggest that our earth also acts like a magnet? Yes, it does.
3. Unlike poles of two magnets attract each other and like poles repel (Fig. 20.1).
4. The poles of a magnet are inseparable, i.e. the simplest specimen providing magnetic field is a magnetic dipole.
5. When a magnet is brought close to a piece of iron, the nearer end of the piece of iron acquires opposite polarity and the farther end acquires same polarity. This phenomenon is called magnetic induction.

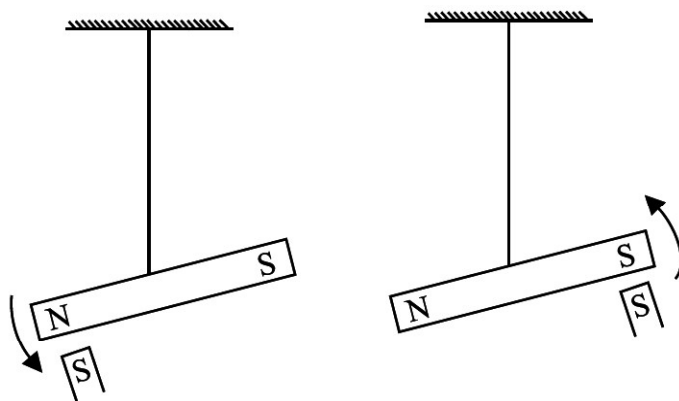


Fig. 20.1 : Unlike poles of two magnets attract each other and like poles repel.

## 20.2) MAGNETIC FIELD AND FIELD LINES

Interactions between magnets or a magnet and a piece of iron essentially represent action at a distance. This can be understood in terms of magnetic field. The region around a magnet in which a magnetic needle (another magnet) experiences a torque (force) and rests in a definite direction, is called 'magnetic field'. This field is represented by a vector quantity  $\mathbf{B}$ , called magnetic field or magnetic induction.

The representation of magnetic fields by magnetic field lines is very useful in visualizing the strength and direction of the magnetic field. Magnetic field lines are imaginary lines along which a magnetic north pole would move. The magnetic field can be represented by the field lines in the following way:

- The direction of magnetic field vector  $\mathbf{B}$  at any point is given by the tangent to the field line at that point.
- The number of field lines that pass through unit area of a surface held perpendicular to the lines is proportional to the strength of magnetic field ( $\mathbf{B}$ ) in that region. Thus, the magnetic field  $\mathbf{B}$  is large where the field lines are closer together and smaller where they are far apart.

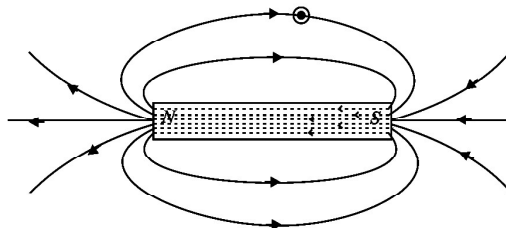


Fig. 20.2 : Magnetic field lines of a bar magnet

- Outside the magnet, the field lines run from north pole to south pole and inside it, these run from south pole to north pole forming closed curves (see Fig. 20.2).
- Two magnetic field lines can never cross each other.

Fig. 20.3 shows a certain number of field lines passing through parallel surfaces  $S_1$  and  $S_2$ . The surface area of  $S_1$  is same as that of  $S_2$  but the number of field lines passing through  $S_1$  is greater than those passing through  $S_2$ . Hence, the number of lines per unit area passing through  $S_1$  is greater than that through  $S_2$ . We can, therefore, say that the magnetic field in the region around P is stronger than that around Q.

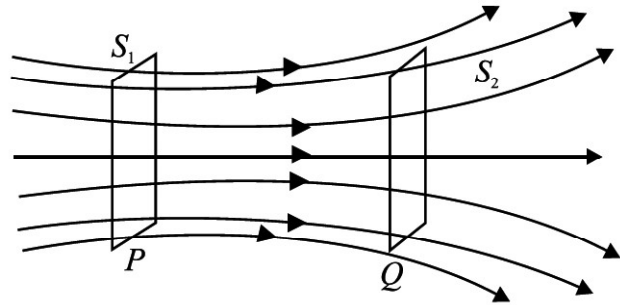


Fig. 20.3 : Magnetic field lines passing through two parallel surfaces

### Intext Questions 20.1

1. You are given a magnet. How will you locate its north pole?
2. You are provided two identical looking iron bars. One of these is a magnet. Using just these two, how will you identify which of the two is a magnet.
3. You are given a thread and two bar magnets. Describe a method by which you can identify the polarities of the two magnets.

## 20.3 THE DIPOLE IN A UNIFORM MAGNETIC FIELD

A magnetic dipole is made up of two poles that are diametrically opposed and separated by a small distance. A small bar magnet can be considered as a magnetic dipole.

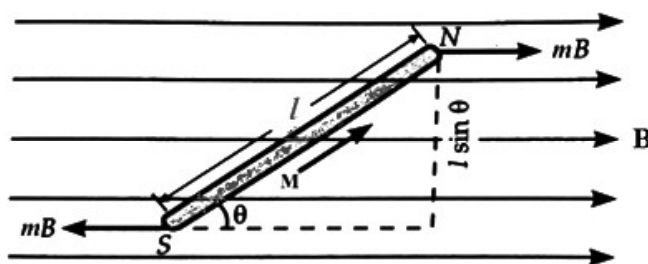


Fig. 20.4 : A small bar magnet suspended in a uniform magnetic field of the strength B.

Let us place a small bar magnet (N-S) of length  $l$  and pole strength of  $m$  in a uniform magnetic field of induction  $B$  at an angle  $\theta$  to the direction of the magnetic field. The force of magnetic field ( $mB$ ) on the north pole of the bar magnet acts in the direction of the magnetic field, while the force ( $mB$ ) on the south pole acts in the opposite direction of the magnetic field. These two new forces are the same and the opposite of each other. As a result, a couple or a restoring torque ( $\tau$ ) is created and hence we have,



Restoring torque = (Normal distance between the forces)  $\times$  (Force)

Since from the Fig. 20.4, normal distance between the forces acting on north and south poles is  $l \sin \theta$ . We can write,

$$\begin{aligned}\tau &= (l \sin \theta) \times (m B) \\ \tau &= (ml) B \sin \theta\end{aligned}\quad (20.1)$$

Here, the product of the pole strength and length of a bar magnet is defined as the magnetic moment of that bar magnet and it is denoted by  $M (= ml)$ . Hence, we can write the above equation as,

$$\tau = MB \sin \theta \quad (20.2)$$

In the vector form,

$$\vec{\tau} = \vec{M} \times \vec{B} \quad (20.3)$$

This is known as restoring torque and due to this restoring torque, a potential energy is stored in the magnet. An expression for magnetic potential energy can be obtained as

$$\begin{aligned}U_m &= \int MB \sin \theta \, d\theta \\ U_m &= -MB \cos \theta\end{aligned}\quad (20.4)$$

Eqn. 20.4 shows that the potential energy is minimum ( $= -MB$ ) at  $\theta = 0^\circ$  (most stable position) and maximum ( $= +MB$ ) at  $\theta = 180^\circ$  (most unstable position).

## 20.4) MAGNETIC FLUX ( $\phi_B$ ) AND MAGNETIC FIELD INDUCTION ( $B$ )

The magnetic field lines passing through a surface area in a magnetic field defined as the magnetic flux ( $\phi_B$ ) through that surface area. It is a scalar quantity. The SI unit of magnetic flux is 'Weber'.

The magnetic induction can also be defined in terms of magnetic field lines. The magnetic field lines passing through a unit surface area placed normal to the field lines measures the magnitude of magnetic induction or magnetic flux density  $B$ . The SI unit of magnetic induction is weber / meter<sup>2</sup> (or) Tesla.

Obviously, in a region, the field lines are spacing close to each, the greater is the magnetic induction. The tangent to the field line at any point gives the direction of magnetic field induction ( $B$ ) at that point. The magnetic field lines simply represent graphically how magnetic field induction ( $B$ ) varies (greater at the poles) throughout a certain region of space. Here the surface area is considered as vector, known as the surface vector  $\vec{A}$  and its direction is being along the normal drawn in the outward direction from the surface as shown in Fig. 20.5.

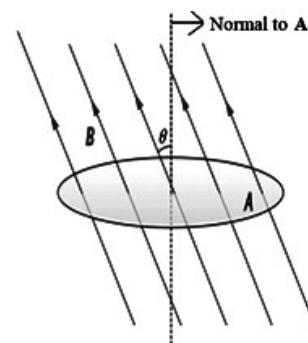


Fig. 20.5 : Magnetic flux through a surface area A in a magnetic field of induction B.

Mathematically, magnetic flux ' $\phi_B$ ' through a surface area 'A', is a magnetic field of induction 'B' is defined as,

$$\phi_B = \vec{B} \cdot \vec{A} = BA \cos \theta \quad (20.5)$$

Where  $\theta$  is the angle between magnetic induction vector  $\vec{B}$  and area vector  $\vec{A}$ .

## 20.5) MAGNETISM AND GAUSS'S LAW

Consider a small vector area element  $\vec{\Delta A}$  of a closed surface as shown in Fig. 20.6. Let  $\Delta\phi_B$  be the magnetic flux through  $\vec{\Delta A}$ . Then it is defined as  $\Delta\phi_B = \vec{B} \cdot \vec{\Delta A}$ , where  $\vec{B}$  is the magnetic field at  $\vec{\Delta A}$ . As the closed surface area A is divide into small area elements like  $\vec{\Delta A}$ , the total flux through the closed surface area A can be calculated by adding the flux through all the area elements over the closed surface area A. Hence, the net flux  $\phi_B$  is,

$$\phi_B = \sum_{\text{All area elements}} \vec{B} \cdot \vec{\Delta A} \quad (20.6)$$

Since the magnetic field lines are continuous loops, all closed surfaces have as many magnetic field lines going in as coming out. Hence, the net magnetic flux through a closed surface is zero. Hence, we can write,

$$\phi_B = \sum_{\text{All area elements}} \vec{B} \cdot \vec{\Delta A} = 0 \quad (20.7)$$

This is the expression for gauss's law in magnetism. Gauss's law for magnetism states that the net magnetic flux through any closed surface is zero.

The difference between the Gauss's law of magnetism and that for electrostatics is a reflection of the fact that isolated magnetic poles (also called monopoles) does not exist. There are no sources or sinks of magnetic field (B) and the simplest magnetic element is a dipole or a current loop. All magnetic phenomena can be explained in terms of an arrangement of dipoles and/or current loops.

## 20.6) MAGNETISATION AND MAGNETIC INTENSITY

When a piece of iron is subjected to a magnetic field, it becomes magnetized, and the level of its magnetism depends on the strength of the field. We therefore need a quantity to describe the degree to which a body is magnetized.

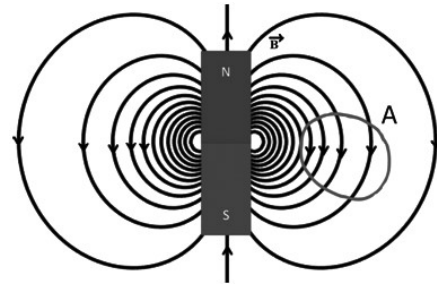
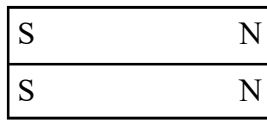


Fig. 20.6 : The field lines of a bar magnet and a closed surface area A.



(a): Two similar magnets are placed side by side to form a compound magnet,



(b) Two similar magnets are placed end to end to form a compound magnet

**Fig. 20.7**

Consider two bar magnets of the same size and shape, each having the same pole strength  $m$  and length  $l$ . If these are placed side by side as shown in the Fig. 20.7 (a), then the poles strength is added and the magnetic moment [ $M = (2m)l = 2(ml)$ ] becomes double to the moment of the individual magnet. If the two magnets are placed end to end as shown in Fig. 20.7 (b), the adjacent poles cancel and length will be added. Then the magnetic moment  $M = m(2l) = 2(ml)$  becomes double to that of the individual magnet. Evidently, the total magnetic moment is the sum of the magnetic moments of the individual magnets.

In these examples, the magnetic moment was doubled by doubling the volume. But, the magnetic moment per unit volume has not changed and it is therefore a quantity that describes the degree to which the magnets are magnetized. It is called the intensity of magnetization ( $I$ ), or simply it is also known as the magnetization. Hence, we can write the expression,

$$\text{Intensity of Magnetisation} = \frac{\text{Total magnetic moment}}{\text{Volume}}$$

$$I = \frac{M}{V} \quad (20.8)$$

Where  $V$  is the volume and  $M$  is the total magnetic moment which can be known as  $M = ml$ .

From the above equation,

$$I = \frac{ml}{V} = \frac{ml}{A \times l} = \frac{m}{A} \quad (20.9)$$

Where  $A$  is the cross-sectional area of the magnet. Therefore, we can also define the magnetisation ( $I$ ) as the pole strength per unit area of cross section.

Since the CGS unit of magnetic moment ( $M$ ) is  $\frac{\text{erg}}{\text{oersted}}$ , the unit of magnetization ( $I$ ) is  $\frac{\text{erg}}{\text{oersted} - \text{cm}^3}$  whereas the SI unit of magnetisation is  $\frac{\text{Ampere}}{\text{meter}}$ .

To describe the magnetic properties of materials, following some other terms are also required. In the magnetisation of a material, magnetic fields are applied. When such magnetic fields pass through magnetic materials which themselves contribute internal magnetic fields. Here an ambiguity can arise about what part of the field comes from the external field and

what comes from the material itself. It has been common practice to define another magnetic field quantity, usually called the “magnetic field intensity or magnetising force” designated by  $H$ . It can be defined by the relationship,

$$H = \frac{B}{\mu} = \frac{B}{\mu_0} - M \quad (20.10)$$

Where  $\mu$  is the permeability of the medium (material) and  $\mu_0$  is the permeability of vacuum or free space. Magnetic field intensity  $H$  is the magnetic influence from external field in the material and independent of the material’s magnetic response. The relationship for  $B$  can be written in the equivalent form as,

$$B = \mu_0 (H + M) \quad (20.11)$$

$H$  and  $M$  will have the same units,  $\frac{\text{Ampere}}{\text{meter}}$ . The quantity  $M$  in this relationship is called the magnetization of the material. To further distinguish  $B$  from  $H$ , the magnetic field intensity  $H$  might be a thought of external magnetic field whereas the magnetic field induction  $B$  is the total magnetic field including the contribution made by the magnetic properties of the materials in the field.

In the Eqn. 20.10, the quantity  $\mu$  is known as magnetic permeability. It is defined as the ability of a material to permit the passage of magnetic field lines through it.

$$\text{Magnetic permeability } (\mu) = \frac{B}{H} \quad (20.12)$$

The ratio of the intensity of magnetisation ( $I$ ) induced in the material to the magnetising force ( $H$ ) applied, is called magnetic susceptibility and it is denoted by  $\chi$ .

$$\chi = \frac{I}{H} \quad (20.13)$$

From the above Eqns. 20.12 and 20.13, relation between magnetic permeability and susceptibility is given by

$$\mu = \mu_0 (1 + \chi) \quad (20.14)$$

Where  $\mu_0$  is known as the permeability of free space. The ratio of permeability of a media (material) to the permeability of free space is defined as the relative permeability ( $\mu_r$ ).

$$\text{Relative permeability } (\mu_r) = \frac{\mu}{\mu_0} \quad (20.15)$$

## 20.7) MAGNETIC PROPERTIES OF MATERIALS

Based on the behaviour of materials in magnetic field, we can divide them broadly into three categories: (i) Diamagnetic materials are feebly repelled by a magnet. (ii) Paramagnetic

materials are feebly attracted by a magnet. (iii) Ferromagnetic materials are very strongly attracted by a magnet. Substances like iron, nickel and cobalt are ferromagnetic. Let us study magnetic behaviour of materials in some details.

### 20.7.1 Diamagnetic materials

Diamagnetic materials are weakly magnetised materials in a direction opposite to the applied magnetic field. Examples are Gold, silver, lead, copper, antimony, bismuth, silicon, mercury, water, air, hydrogen, nitrogen etc.

In general, the magnetic moments of atoms are oriented such that the vector sum of the magnetic moments of atoms is zero. When such a diamagnetic material is placed in an external magnetic field, the material as a whole acquires a net magnetic moment in a direction opposite to the applied field. Atoms of diamagnetic materials do not possess any unpaired electrons.

#### Characteristics:

1. Diamagnetic materials possess small and negative susceptibility. Hence, they have  $\chi_m < 0$ .
2. The magnetic moment of the atoms of diamagnetic material aligns in the opposite direction to applied magnetic field.
3. The atomic magnets of these materials are slightly repelled by a magnetic field.
4. When diamagnetic materials placed in a non-uniform magnetic field, they tend to move from the stronger to the weaker part of the field.
5. They do not retain the magnetic properties when the external field is removed.
6. Susceptibility is independent of temperature of the material.

### 20.7.2 Paramagnetic materials

Paramagnetic materials are weakly magnetised materials in the same direction as the applied magnetic field. Examples are aluminium, platinum, manganese, chromium, sodium, calcium, lithium, tungsten, niobium, oxygen etc.

In the paramagnetic materials, the orbital and the spin magnetic moments of atoms are oriented in such a way that, each atom has a permanent magnetic dipole moment. However, due to thermal motion (vibration), magnetic moments have random orientations. As a result, the net magnetic moment is zero. When these materials are placed in an external magnetic field, each atomic magnet tends to align in the direction of the field. The materials acquire a net magnetic moment and get magnetised. The atoms of paramagnetic materials possess a few unpaired electrons.

#### Characteristics:

1. Paramagnetic materials possess small and positive susceptibility. Hence, they have  $\chi_m > 0$ .
2. The magnetic moment of the atoms of paramagnetic materials aligns in the same

direction as that of the applied magnetic field.

3. The atomic magnets of these materials are slightly attracted by a magnetic field.
4. When paramagnetic materials are placed in a magnetic field, they tend to move from weaker to the stronger parts of the non-uniform external magnetic field
5. These materials do not retain the magnetic properties when the external field is removed.
6. Susceptibility of paramagnetic materials depend on the temperature of the material.

### 20.7.3 Ferromagnetic materials

Ferromagnetic materials are strongly magnetised materials in the same direction as that of applied magnetic field and retains its magnetic moment even after the removal of the applied field. Examples are Iron, cobalt, nickel etc.

In ferromagnetic materials, the contribution of the spin magnetic moment to the magnetic dipole moment of atoms is very large. A ferromagnetic substance contains small regions called domains. All magnetic dipoles in a domain are fully aligned. The magnetization of domains is maximum. But the domains are randomly oriented. As a result, the total magnetic moment of the sample is zero. When we apply an external magnetic field, the domains slightly rotate and align themselves in the direction of the field giving rise to resultant magnetic moment. The process can be easily understood with the help of a simple diagram shown in Fig. 20.8.

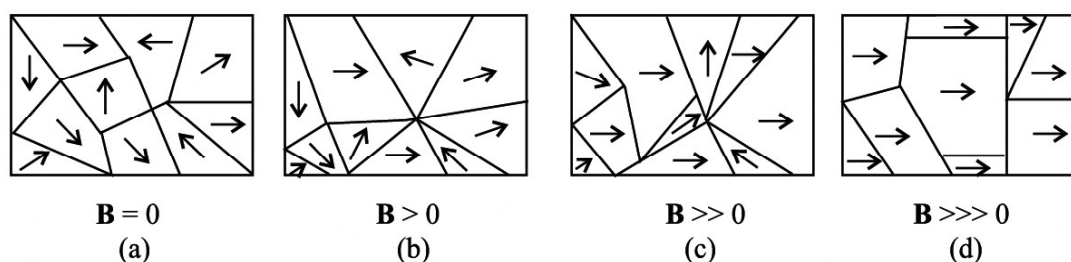


Fig. 20.8 : Domains in a ferromagnetic substance.

Fig. 20.8 (a) shows ten domains. For simplicity we take a two-dimensional example. All the domains are so directed that the total magnetization of the sample is zero. Fig. 20.8 (b) shows the state after the application of an external magnetic field. The boundaries of the domains (Domain Walls) reorganise in such a way that the size of the domain having magnetic moment in the direction of the field becomes larger at the cost of others. On increasing the strength of external field, the size of favourable domains increases, and the orientation of the domain changes slightly resulting in greater magnetization [Fig. 20.8 (c)]. Under the action of very strong applied field, almost the entire volume behaves like a single domain giving rise to saturated magnetization as shown in the Fig. 20.8 (d). When the external field is removed, the sample retains net magnetization. The domain in ferromagnetic samples can be easily seen with the help of high-power microscope.

When the temperature of a ferromagnetic substance is raised beyond a certain critical value, the substance becomes paramagnetic. This critical temperature is known as Curie temperature  $T_c$ .

**Characteristics:**

1. Ferromagnetic materials possess large and positive susceptibility and the susceptibility is  $\chi_m > 1$ .
2. The magnetic moment of the atoms of ferromagnetic materials aligns in the same direction as that of applied magnetic field.
3. The atomic magnets of the ferromagnetic material have strong attraction towards magnetic fields.
4. When ferromagnetic materials are placed in a magnetic field, they have a strong tendency to move from weaker to the stronger parts of the non-uniform external magnetic field.
5. Ferromagnetic materials retain their magnetic properties even after the external field is removed.
6. Susceptibility of ferromagnetic materials depends on the temperature of the material.
7. In ferromagnetic materials, the magnetization is not proportional to the applied field and exhibits a property called Hysteresis.

**20.8) MAGNETIC FIELD OF THE EARTH**

The directive property of magnets could be explained by considering that the earth acts as a magnet, i.e., as if a large bar magnet is placed inside the earth. The south pole of this magnet is considered near the geographical north pole and the magnetic north pole near the geographical south pole as shown in the Fig. 20.9. To explain the elements of the earth's magnetic field, consider  $RR_1$  the rotational axis of earth and  $MM_1$  is the magnetic axis of the earth.

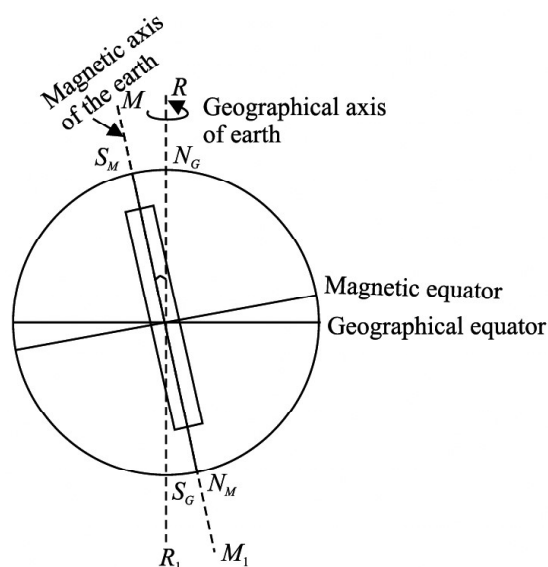


Fig. 20.9 : Magnetic field of the Earth

**Activity 20.1**

Let us perform a thought experiment with a magnetic needle. (You can actually perform the experiment with a globe containing a bar magnet along its axis of rotation with north pole of the magnet pointing south.) Suspend the needle freely in such a manner that it can rotate in horizontal as well as vertical planes. If the needle is near the equator on earth's surface, it rests in horizontal plane. Suppose this needle is taken to places in the northern hemisphere, the needle rotates in the vertical plane and the north pole dips towards the earth as we move towards geographical north pole. Finally at a point very near to Hudson Bay in Canada, the north pole of the needle will point vertically downward. This place, located at  $6^\circ$  east of north, is considered to be the

south pole of the earth's magnet. This place is about 650 km away from the earth's geographical north pole. If we take the same magnetic needle to places in the southern hemisphere, the south pole of the needle will dip downward and point vertically downward at a point 650 km west of the geographical south. This point could be considered as the north pole of the earth's magnet. From this, we conclude that the magnetic axis of the earth does not coincide with the geographical axis.

An important aspect of earth's magnetic field is that it does not remain constant. Its magnitude and direction change with time.

### Elements of the Earth's Magnetic Field

Three measurable quantities are used to describe the magnetic field of earth. These are called elements of earth's magnetic field:

- Inclination or dip ( $\delta$ );
- Declination ( $\theta$ ) and
- Horizontal component of the earth's field ( $B_H$ ).

#### 20.8.1 Inclination or Dip ( $\delta$ )

If you suspend a magnetic needle freely at a place, you will observe that the needle does not rest in the horizontal plane. It will point in the direction of the resultant intensity of earth's magnetic field.

Fig. 20.10 shows the plane PCDE, which is the magnetic meridian at the point P (i.e., the vertical plane passing through the north and south poles of the earth's magnet) on the surface of the earth and PABC is the geographic meridian (i.e., the vertical plane passing through the geographical north and south poles of the earth). Suppose that PR represents the magnitude and direction of the earth's magnetic field at the point P. Note that PR makes an angle  $\delta$  with the horizontal direction. This angle is known as inclination or dip at P on the surface of the earth.

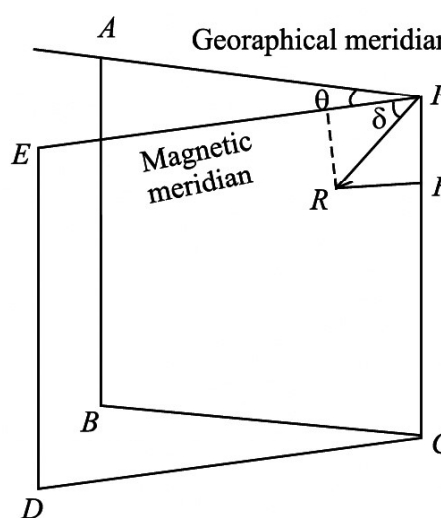


Fig. 20.10 : Elements of earth's magnetic field.

The angle which the earth's magnetic field makes with the horizontal direction in the magnetic meridian is called the dip or inclination.

#### 20.8.2 Declination ( $\theta$ )

Refer to Fig. 20.10 again. The plane PCDE contains the magnetic field vector (PR) of the earth. The angle between the planes PCDE and PABC is called the declination at the point P. It is shown as angle  $\theta$ .



The angle which the magnetic meridian at a place makes with the geographical meridian is called the declination at that place.

### 20.8.3 Horizontal component ( $B_H$ )

Fig. 20.10 shows that PR is the resultant magnetic field at the point P. PE represents the horizontal component and PF the vertical component of the earth's magnetic field in magnitude and direction. Let the magnetic field at the point P be B. The horizontal component of the earth magnetic field is

$$B_H = B \cos \delta \quad (20.16)$$

And the vertical component is

$$B_V = B \sin \delta \quad (20.17)$$

By squaring and adding Eqn. 20.16 and Eqn. 20.17, we have

$$B_H^2 + B_V^2 = B^2 \cos^2 \delta + B^2 \sin^2 \delta = B^2$$

On dividing Eqn. 20.17 by Eqn. 20.16, we have

$$\frac{B_V}{B_H} = \tan \delta \quad (20.18)$$

## WHAT YOU HAVE LEARNT

- The property of any object by virtue of which it can attract a piece of iron is called magnetism.
- Every magnet has two poles. These are inseparable.
- Unlike poles of two magnets attract each other and like poles repel.
- The region around a magnet in which a magnetic needle (another magnet) experiences a torque (force) is called 'magnetic field'.
- Magnetic field lines are imaginary lines along which a magnetic north pole would move in a magnetic field.
- When a bar magnet or magnetic dipole of moment M is suspended in an magnetic field of induction B at an angle  $\theta$  to the direction of the magnetic field, a couple or a restoring torque  $\tau = MB \sin \theta$  is created on the magnet or dipole.
- The magnetic field lines passing through a surface area in a magnetic field defined as the magnetic flux ( $\phi_B$ ) through that surface area.
- The SI unit of magnetic flux is 'weber'.
- The magnetic field lines passing through a unit surface area placed normal to the field lines measures the magnitude of magnetic induction or magnetic flux density B.
- The SI unit of magnetic induction is weber / meter<sup>2</sup> (or) Tesla.

- Gauss's law for magnetism states that the net magnetic flux through any closed surface is zero.
- The magnetic moment induced in unit volume of a magnetic material when it is subjected to an external magnetic field, is defined as the magnetisation ( $I$ ).
- Materials are broadly classified into three categories based on their magnetic behaviour and they are: (i) Diamagnetic materials (ii) Paramagnetic materials and (iii) Ferromagnetic materials.
- The directive property of magnets could be explained by considering that the earth acts as a magnet. Hence earth is a big magnet and around the earth there is field known as the earth's magnetic field.
- The angle which the earth's magnetic field makes with the horizontal direction in the magnetic meridian is called the dip or inclination ( $\delta$ ).
- The angle which the magnetic meridian at a place makes with the geographical meridian is called the declination ( $\theta$ ) at that place.
- If  $B$  is the resultant vector of the earth's magnetic field, the horizontal component is  $B_H = B \cos \delta$ .

### TERMINAL EXERCISE

1. Define the magnetic field and magnetic field lines.
2. Define magnetic flux, magnetic field induction and mention their units.
3. State and explain the Gauss's law in magnetism.
4. Define magnetisation and magnetic intensity.
5. Classify the magnetic materials and explain them.
6. What are the elements of earth magnetic field? Explain them.
7. A small piece of the material is brought near a magnet. Complete the following Table by filling up the blanks by writing Yes or No.

Material	Repulsion		Attraction	
	Weak	Strong	Weak	Strong
Diamagnetic				
Paramagnetic				
Ferromagnetic				

8. You have to keep two identical bar magnets packed together in a box. How will you pack and why?

N    S    OR    N    S  
 N    S                    S    N

## TOSS

- The magnetic force between two poles is 80 units. The separation between the poles is doubled. What is the force between them?
- The length of a bar magnet is 10 cm and the area of cross-section is  $1.0 \text{ cm}^2$ . The magnetization  $I = 10^2 \text{ A m}^{-1}$ . Calculate the pole strength.
- Two identical bar magnets are placed on the same line end to end with north pole facing north pole. Draw the lines of force, if no other field is present.
- The points, where the magnetic field of a magnet is equal and opposite to the horizontal component of magnetic field of the earth, are called neutral points,
  - Locate the neutral points when the bar magnet is placed in magnetic meridian with north pole pointing north.
  - Locate the neutral points when a bar magnet is placed in magnetic meridian with north pole pointing south.
- If a bar magnet of length 10 cm is cut into two equal pieces each of length 5 cm. Then what is the pole strength of the new bar magnet compare to that of the old one.
- A 10 cm long bar magnet has a pole strength 10 A m. Calculate the magnetic field at a point on the axis at a distance of 30 cm from the centre of the bar magnet.

## ANSWERS TO INTEXT QUESTIONS

### 20.1

- Suspend the magnet with a thread at its centre of mass. Let it come to equilibrium. The end of the magnet which points towards geographical north is its north pole.
- Bring the ends of any two bars closer together. If there is attraction between them, one of the bars is a magnet and the other is an iron bar. Now lay down one of these bars on the table and stick along its length with the other. If uniform force is experienced, the bar in hand is a magnet and that on the table is iron piece. If non-uniform force is experienced, reverse is the case.
- Suspending one of the bar magnets with thread, we can find its south pole. Then the end of the second magnet, which is repelled by the first, is its south pole.

## ANSWERS TO TERMINAL EXERCISE

- 20 units
- $1.02 \times 10^{-2} \text{ Am}$
- Same. Pole strength of the magnet does not depend on the length of magnet.
- $2.3 \times 10^{-6} \text{ T}$



# MOVING CHARGES AND MAGNETISM

## INTRODUCTION

In the previous chapter, we learnt the electric field and magnetic fields. Did you ever think of some relation between electricity and magnetism? Such a relationship was discovered by Oersted in 1820. Now we know, for sure, how intimately magnetism and electricity are related.

In this lesson, you will learn the magnetic effects of electric current. The behaviour of current carrying conductors and moving charges in a magnetic field are also discussed. On the basis of these principles, we will discuss the working of electric devices like motors and measuring devices like an ammeter, a voltmeter and a galvanometer.

## OBJECTIVES

After studying this lesson, you should be able to

- describe the magnetic effect of electric current: Oersted's experiment;
- state Biot-Savart's law and explain its applications;
- derive the expression for magnetic field induction due to electric current;
- indicate the direction of magnetic field due to current in a conductor;
- explain Ampere's circuital law and its application;
- describe the motion of a charged particle in a uniform magnetic field;
- derive an expression for the force experienced by a current carrying conductor placed in a uniform magnetic field;
- derive an expression for the force between two infinitely long current carrying conductors placed parallel to each other;
- explain the working principle of a galvanometer, an ammeter and a voltmeter.

## 21.1) ELECTRICITY AND MAGNETISM : BASIC CONCEPTS

You now know that flow of electrons in a conductor due to a potential difference across it constitutes electric current. The current flowing in a conductor is seen to exert a force on a free magnetic needle placed in a region around it. A magnetic needle is also affected by a magnet and hence we say that a current carrying conductor has a magnetic field around it. The magnetic field  $B$  is visualized by magnetic field lines. You will learn about these and some more terms such as magnetic permeability later in this lesson.

## 21.2) MAGNETIC FIELD AROUND AN ELECTRIC CURRENT

Let us do a simple experiment.

### Activity 21.1

Take a 1.5 Volt battery, a wire about 1 m in length, a compass needle and a match box. Wind 10 – 15 turns of the electric wire on its base. Under the windings, place a compass needle, as shown in Fig. 21.1. Place the match box on the table so as to have the wires running along the north – south direction. Connect the free ends of the wire to the battery. What happens to the needle? You will observe that needle shows deflection. This means that there is a magnetic field in and around the coil.

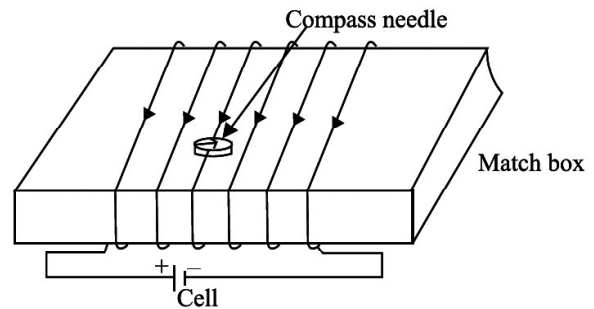
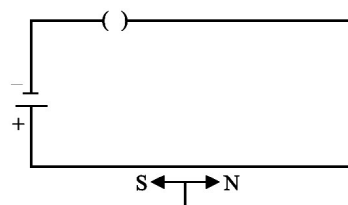
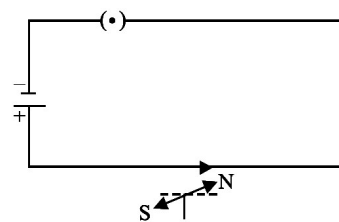


Fig. 21.1 : Demonstration of magnetic field due to electric current.

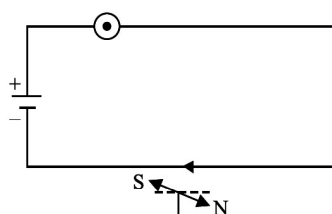
The deflection will reverse if you reverse the direction of current by changing the terminals of the battery. When there is no current in the wire, the compass needle placed near the current carrying wire as shown in the Fig. 21.2 (a) points in the north – south direction. When the current is flowing through the wire compass needle gets the reflection as shown in Fig. 21.2 (b) and if the current direction is reversed, then compass needle is also reversed as shown in Fig. 21.2 (c). When a magnetic needle is brought close to a vertical current carrying wire, the magnetic field lines are found as concentric circles around the wire, as shown in Fig 21.2 (d).



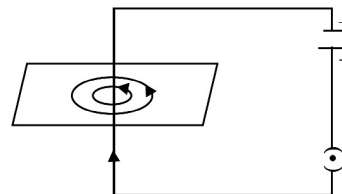
(a) No current, No deflection



(b) Current towards north, deflection of north pole towards west



(c) When direction of current is reversed, direction of deflection is reversed



(d) Circular field lines around a straight current carrying conductor

Fig. 21.2 : Magnetic field around a current carrying conductor

In 1820, Hans Christian Oersted, Professor of Physics at Copenhagen in Denmark performed similar experiments and established that there is a magnetic field around a current carrying conductor.

### 21.3) BIOT-SAVART'S LAW

Biot-Savart's law gives a quantitative relationship between current in conductor and the resulting magnetic field at a point in the space around it. Each part of a current carrying conductor contributes to magnetic field around it. The net value of  $B$  at a point is thus the combined effect of all the individual parts of the conductor. As shown in Fig. 21.3, the net magnetic field due to any current carrying conductor is the vector sum of the contributions due to the current in each infinitesimal element of length  $\Delta l$ .

Experiments show that the field  $\Delta B_0$  due to an element  $\Delta l$  is

- 1) proportional to the current flowing through the conductor,  $I$
- 2) proportional to length of the element  $\Delta l$
- 3) proportional to the angle between the element  $\Delta l$  and the line joining the element to the observation point  $P$  and
- 4) inversely proportional to the square of the distance ( $r$ ) of observation point  $P$  from the element  $\Delta l$

Thus, we can write,

$$|\Delta B_0| \propto \frac{I \Delta l \sin \theta}{r^2}$$

$$|\Delta B_0| = \frac{\mu_0}{4\pi} \frac{I \Delta l \sin \theta}{r^2} \quad (21.1)$$

where  $\mu_0$  is permeability of vacuum. Its value is  $4\pi \times 10^{-7} \text{ Wb A}^{-1} \text{ m}^{-1}$ . The value of permeability of air is also nearly equal to  $\mu_0$ . If the conductor is placed in a medium other than air, the value of the field is altered and is given by  $B = \mu B_0$ . Here  $\mu$  represents the permeability of the medium.

**Direction of B :** Magnetic field at a point is a vector quantity. The direction of  $B$  may be determined by applying the right-hand grip rule. To apply this rule, let us consider the direction of the field produced in some simple cases. As shown in the Fig. 21.4 (a), grasp the wire in your right hand so that the thumb points in the direction of the current. Then

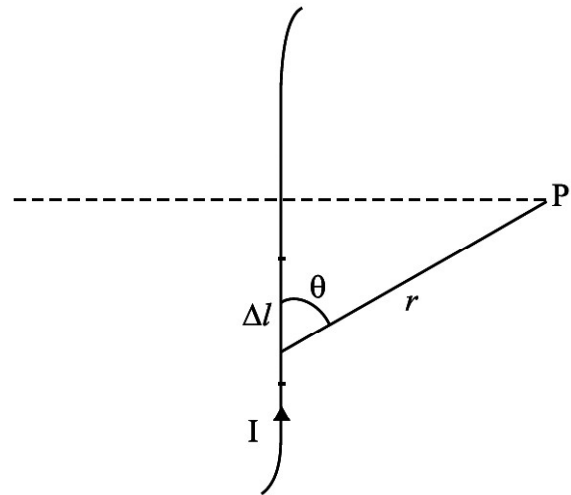
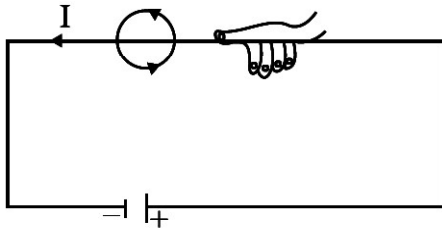
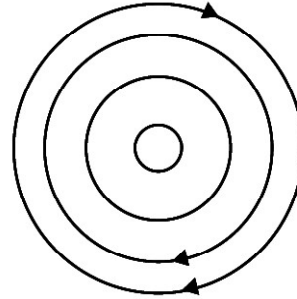


Fig. 21.3 : Magnetic field at  $P$  due to a current element  $\Delta l$ .

the curled fingers of the hand will point in the direction of the magnetic field. To represent the magnetic field on paper, let us consider that current is flowing in to the plane of the paper. Then according to the right-hand rule, the field lines shall be in the plane of the paper as shown in the Fig. 21.4 (b).



(a) Right hand rule: thumb in the direction of current, field lines in the direction of curling fingers



(b) when current is in to the plane of paper, the field lines shall be in the plane of paper, according to the right-hand rule

**Fig. 21.4 :** Direction of magnetic field

### 21.3.1 Applications of Biot-savart's law: magnetic field on the axis of a circular current loop

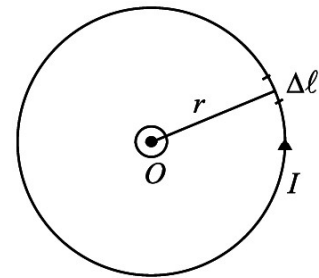
You now know that the Biot-Savart's law gives the magnitude of the magnetic field. Let us now apply it to find the field around conductors of different shapes. Note that to calculate the net field due to different segments of the conductor, we have to add up the field contributions due to each one of them. We first consider a circular coil carrying current and calculate magnetic field at its centre.

#### Magnetic field at the centre of a circular coil carrying current

Refer to Fig. 21.5. It shows a circular coil of radius  $r$  carrying current  $I$ . To calculate magnetic field at its centre  $O$ , we first consider a small current element  $\Delta l$  of the circular coil. Note that the angle between current element  $\Delta l$  and  $r$  is  $90^\circ$ . From Eqn. 21.1, we know that the field at the centre  $O$  due to  $\Delta l$  is

$$|\Delta B| = \frac{\mu_0}{4\pi} I \frac{\Delta l}{r^2} \sin 90^\circ$$

$$|\Delta B| = \frac{\mu_0}{4\pi} I \frac{\Delta l}{r^2} \quad (\text{as } \sin 90^\circ = 1)$$



**Fig. 21.5 :** Circular coil carrying current.

The direction of  $\Delta B$  is normal to the plane of the coil. Since the field due to every element of the circular coil will be in the same direction, the resultant is obtained by adding all the contributions at the centre of the loop. Therefore,

$$|B| = \Sigma|\Delta B| = \frac{\mu_0 I}{4\pi r^2} \Sigma \Delta l = \frac{\mu_0 I}{4\pi r^2} (2\pi r)$$

Hence, magnetic field at the centre of a coil of radius  $r$  carrying current  $I$  is given by

$$|B| = \frac{\mu_0 I}{2r} \quad (21.2)$$

In case, there is more than one loop of wire (say there are  $n$  turns), the field is given by

$$|B| = \frac{\mu_0 nI}{2r} \quad (21.3)$$

You can check the direction of the net field using the rule given in Fig. 21.4. You can use right hand rule in any segment of the coil and will obtain the same result. (Another simple quick rule to identify the direction of magnetic field due to a current carrying coil is the so-called End-rule, illustrated in Fig. 21.6 (a) and (b).

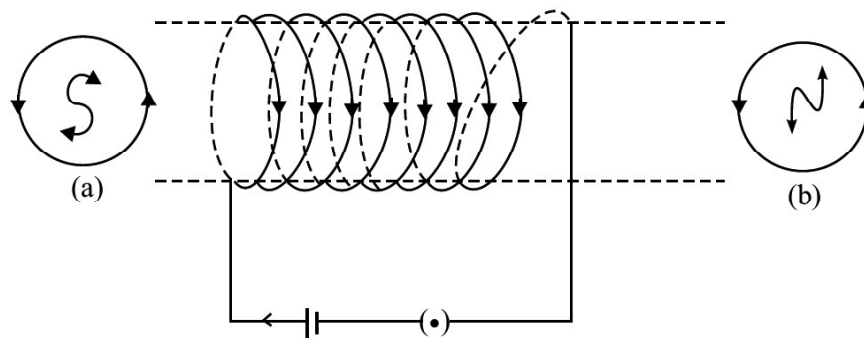


Fig. 21.6 : Direction of magnetic field: End-rule

When an observer looking at the circular coil at its either end, finds the current to be flowing in the clockwise sense, the face of the coil behaves like the south pole of the equivalent magnet, i.e.,  $B$  is directed inwards. On the other hand, if the current is seen to flow in the anticlockwise sense, the face of the coil behaves like the north pole of the equivalent magnet or the field is directed out of that end.

### Intext Questions 21.1

1. What can you say about the field developed by (i) a stationary electron and (ii) a moving electron?
2. Electrons in a conductor are in constant motion due to thermal energy. Why do they not show magnetism till such time that a potential difference is applied across it?
3. A current is flowing in a long wire. It is first shaped as a circular coil of one turn, and then into a coil of two turns of smaller radius. Will the magnetic field at the centre coil change? If so, how much?



### Andre Marie Ampere (1775 – 1836)

French Physicist, mathematician and chemist, Ampere was a child prodigy. He mastered advanced mathematics at the age of 12. A mix of experimental skills and theoretical acumen, Ampere performed rigorous experiments and presented his results in the form of a theory of electrodynamics, which provides mathematical formulation of electricity and its magnetic effects. Unit of current is named in his honour. Lost in his work and ideas, he seldom cared for honours and awards. Once he forgot an invitation by emperor Napoleon to dine with him. His gravestone bears the epitaph: Tendun felix (Happy at last), which suggests that he had to face a very hard and unhappy life. But it never lowered his spirit of creativity



## 21.4) AMPERE'S CIRCUITAL LAW

Ampere's circuital law provides another way of calculating magnetic field around a current carrying conductor in some simple situations.

Consider a closed loop of radius 'r' around a straight wire carrying current 'I' as shown in Fig. 21.7. Then,

Ampere's circuital law states that the line integral of the magnetic field  $B$  around a closed loop is  $\mu_0$  times the total current,  $I$ . Mathematically, we write

$$\oint B \cdot dl = \mu_0 I \quad (21.4)$$

Here, 'dl' is a small length along the closed loop.

Note that this is independent of the size or shape of the closed loop.

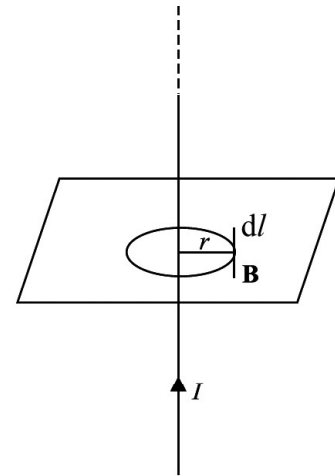


Fig. 21.7 : Ampere's circuital law

### 21.4.1 Applications of ampere's circuital law

We now apply Ampere's circuital law to obtain magnetic field in two simple situations.

- (a) **Magnetic field due to an infinitely long current carrying conductor :** Refer to Fig. 21.8. It shows an infinitely long current carrying conductor POQ carrying current  $I$ . Consider a circular loop of radius  $r$  around it in the plane as shown. Then

$$\Sigma B \cdot dl = B (2\pi r)$$

By applying Ampere's circuital law, we can write

$$\Sigma B \cdot dl = |B| (2\pi r) = \mu_0 I$$

$$\text{or} \quad |B| = \frac{\mu_0 I}{2\pi r} \quad (21.5)$$

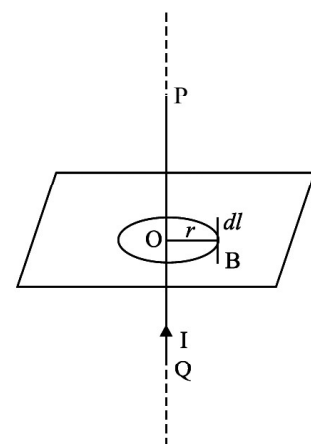


Fig. 21.8 : Infinitely long current carrying conductor

This gives the magnetic field around an infinitely long straight current carrying conductor. Solenoids and toroids are widely used in motors, generators, toys, fan-windings, transformers, electromagnets etc. They are used to provide uniform magnetic field. When we need large fields, soft iron is placed inside the coil.

- (b) **Magnetic field due to a solenoid (The solenoid) :** A solenoid is a coil having a large number of loops set in a straight line with a common axis, as shown in Fig. 21.9. We know that a current  $I$  flowing through a wire, sets up a magnetic field around it. Suppose that the length of the solenoid is  $L$  and it has  $n$  number of turns per unit length, then the enclosed current is  $I_e = I(nL)$ . From Ampere's circuital law

$$B.L = \mu_0 I_e \quad (\text{or})$$

$$B.L = \mu_0 I(nL)$$

$$|B| = \mu_0 nI$$

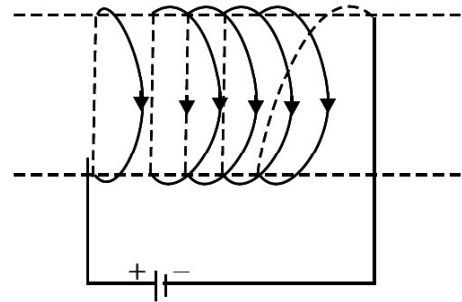


Fig. 21.9 : A solenoid

The direction of the field is along the axis of the solenoid. For long solenoid,  $|B| = \mu_0 nI$  should be correct well inside the solenoid, near its centre. For shorter solenoids the magnitude of  $B$  at the ends is given by

$$|B| = \frac{\mu_0 nI}{2} \quad (21.6)$$

When electric current flows through a solenoid, it generates a magnetic field around it. The magnetic field lines produced by the solenoid resembles the field lines of a bar magnet as shown in the Fig. 21.10.

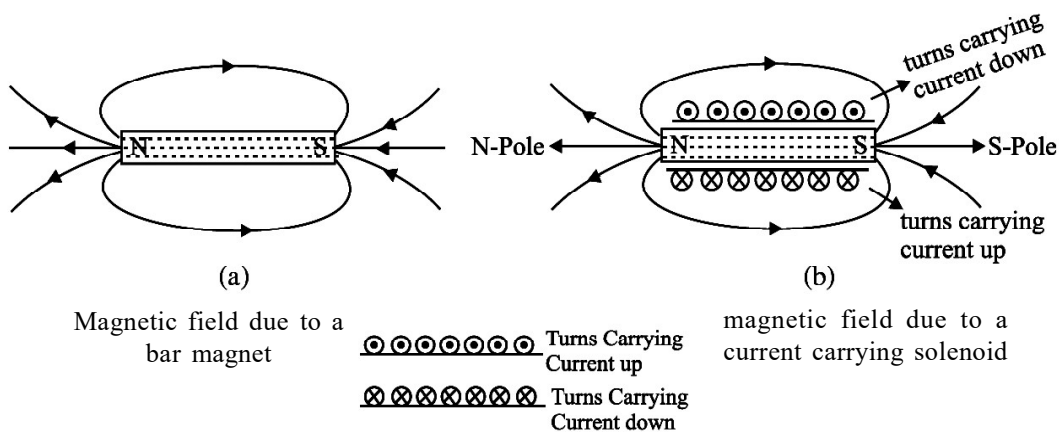


Fig. 21.10 : Solenoid behaves like a bar magnet

### Example 21.1

A 50 cm long solenoid has 3 layers of windings of 250 turns each. The radius of the lowest layer is 2 cm. If the current through it is 4.0 A, calculate the magnitude of  $B$  at (a) near the centre of the solenoid on and about the axis; (b) near the ends on its axis; and (c) outside the solenoid near the middle.

**Solution :**

a) At the centre or near it

$$\begin{aligned}
 |B| &= \mu_0 nI \\
 &= 4\pi \times 10^{-7} \times \frac{3 \times 250}{0.5} \times 4 \\
 &= 16\pi \times 1500 \times 10^{-7} \text{ T} \\
 &= 24\pi \times 10^{-4} \text{ T}
 \end{aligned}$$

b) At the ends

$$\begin{aligned}
 B_{\text{ends}} &= \frac{1}{2} B_{\text{centre}} = \frac{1}{2} (24\pi \times 10^{-4} \text{ T}) \\
 &= 12\pi \times 10^{-4} \text{ T}
 \end{aligned}$$

c) Outside the solenoid the field is zero.

**Example 21.2**

Calculate the distance from a long straight wire carrying a current of 12 A at which the magnetic field will be equal to  $3 \times 10^{-5}$  T.

**Solution :**

$$\begin{aligned}
 |B| &= \frac{\mu_0 I}{2\pi r} \Rightarrow r = \frac{\mu_0 I}{2\pi B} \\
 r &= \frac{4\pi \times 10^{-7} \times 12}{2\pi \times 3 \times 10^{-5}} = 0.25 \text{ m}
 \end{aligned}$$

**Intext Questions 21.2**

- A drawing of the lines of force of a magnetic field provides information on (a) direction of field only (b) magnitude of field only (c) both the direction and magnitude of the field (d) the force of the field.
- What is common between Biot-Savart's law and Ampere's circuital law?
- In the following drawing of lines of force of a non-uniform magnetic field, at which point is the field: (i) uniform, (ii) weakest and (iii) strongest?
- A 10 cm long solenoid is meant to have a magnetic field 0.002 T inside it, when a current of 3 A flows through it. Calculate the required no. of turns.

**Fig. 21.11 :** A typical magnetic field

## 21.5) FORCE ON A MOVING CHARGE IN A MAGNETIC FIELD

When a charged body moves in a magnetic field, it experiences a force. Such a force experienced by a moving charge is called the Lorentz force. The Lorentz force on a particle with a charge  $+q$  moving with a velocity  $v$  in a magnetic field  $\mathbf{B}$  is given by

$$\mathbf{F} = q (\mathbf{v} \times \mathbf{B}) \quad \text{or} \quad |\mathbf{F}| = q v B \sin \theta \quad (21.7)$$

where  $\theta$  is the angle between the directions of  $v$  and  $B$ . The direction of  $F$  is given by Fleming's left-hand rule.

Fleming's left hand rule states that if we stretch the fore finger, the central finger and the thumb of our left hand at right angles to each other and hold them in such a way that the fore finger points in the direction of magnetic field and the central finger points in the direction of motion of positively charged particle, then the thumb will point in the direction of the Lorentz force as shown in the Fig. 21.12.

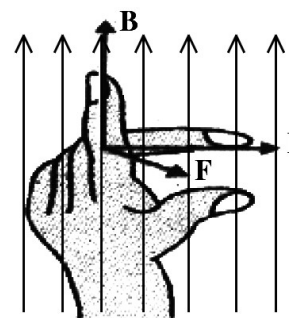


Fig. 21.12 : Fleming's left-hand rule

### Some important points to note

- $F$  is a mechanical force resulting in a pull or a push.
- The direction of force is given by Fleming's left-hand rule.
- In case of negative charges, the central finger should point opposite to the direction of its motion.
- If the charge stops, the force becomes zero instantly.
- Force is zero when charges move along the field  $B$ .
- Force is maximum when charges move perpendicular to the field. Then,  $F = qvB$ .

## 21.6) FORCE ON A CURRENT CARRYING CONDUCTOR IN A UNIFORM MAGNETIC FIELD

The concept of Lorentz force can be easily extended to current carrying conductors placed in uniform magnetic field  $\mathbf{B}$ . Suppose that the magnetic field is parallel to the plane of paper and a conductor of length  $\Delta l$  carrying current  $I$  is placed normal to the field. Suppose further that the current is flowing downward with a drift velocity  $v_d$  and hence each free electron constituting the current experiences a Lorentz force,  $\mathbf{F} = e v_d \mathbf{B}$ .

If there are  $N$  free electrons in the conductor, the net force on it is given by

$$\mathbf{F} = N e v_d \mathbf{B} = (nA \Delta l) e v_d \mathbf{B} \quad (21.8)$$

Where  $n$  denotes the number of free electrons per unit volume. But,  $neAv_d = I$ . Hence,

$$\mathbf{F} = I \Delta l \mathbf{B} \quad (21.9)$$

If the direction of the current makes an angle  $\theta$  with the direction of magnetic field as shown in the Fig. 21.13 (c), then the above Eqn. (21.9) modifies as

$$|\mathbf{F}| = I \Delta l B \sin \theta \quad (21.9 a)$$

If the direction of the current is along the direction of the magnetic field as shown in the Fig. 21.13 (a), then  $\theta = 0$  and force acting on the conductor is zero.

If the direction of the current is perpendicular to the direction of the magnetic field as shown in the Fig. 21.13 (b), then  $\theta = 90^\circ$  and maximum force acts on the conductor i.e.,  $\mathbf{F} = I \Delta l \mathbf{B}$ .

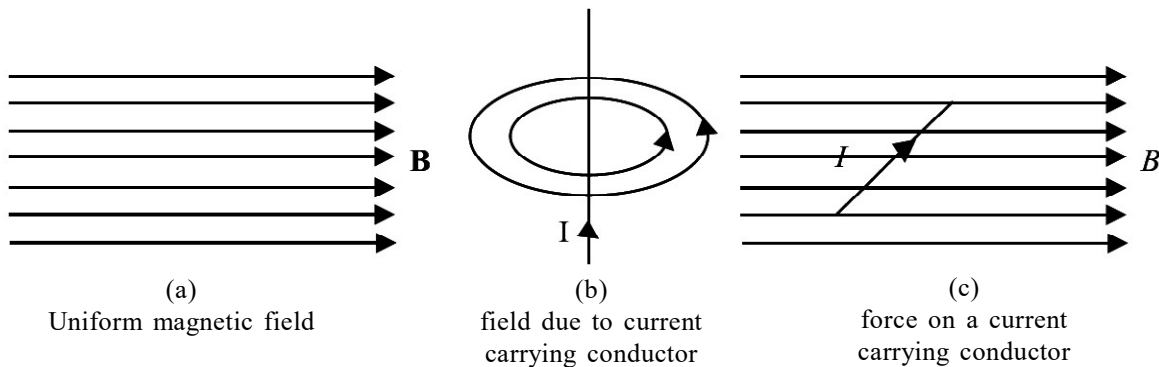


Fig. 21.13

The direction of the force is again given by Fleming’s left-hand rule. Eqn. 21.9 can be used to define the unit of magnetic field in terms of the force experienced by a current carrying conductor. By rearranging terms, we can write

$$\mathbf{B} = \frac{\mathbf{F}}{I \Delta l}$$

Since  $\mathbf{F}$  is taken in Newton,  $I$  in Ampere and  $\Delta l$  in metre, the unit of  $B$  will be  $\text{N A}^{-1} \text{m}^{-1}$ . It is called Tesla (T).

### 21.7) FORCE BETWEEN TWO PARALLEL WIRES CARRYING CURRENT

You now know that every current carrying conductor is surrounded by a magnetic field. It means that it will exert force on a nearby current carrying conductor. The force between two current carrying conductors placed parallel to each other is mutual and magnetic in origin. A current carrying wire has no net electric charge, and hence cannot interact electrically with another such wire.

Fig. 21.14 shows two parallel wires separated by distance  $r$  and carrying currents  $I_1$  and  $I_2$  respectively. The magnetic field due

to one wire at a distance  $r$  from it is  $B_1 = \frac{\mu_0 I_1}{2\pi r}$ .

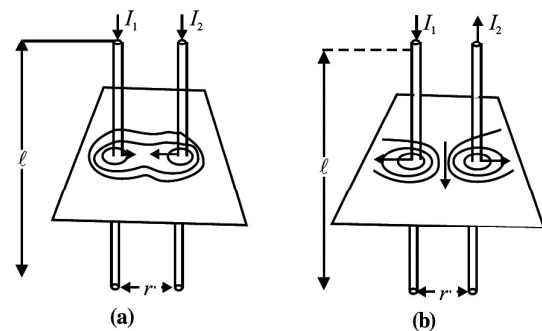


Fig. 21.14 : Experimental demonstration of force between two parallel wires carrying current.

Similarly, the field due to second wire at a distance  $r$  from it will be  $B_2 = \frac{\mu_0 I_2}{2\pi r}$ .

These fields are perpendicular to the length of the wires and therefore the force on a length  $l$  of the other current carrying conductor is given by

$$F = BIl = \frac{\mu_0 I_1}{2\pi r} I_2 l$$

or force per unit length

$$\frac{F}{l} = \frac{\mu_0 I_1 I_2}{2\pi r} \quad (21.10)$$

The forces are attractive when the currents are in the same direction and repulsive while they are in opposite directions. Eqn. 21.10 can be used to define the unit of current.

If  $I_1 = I_2 = 1$  A,  $l = 1$  m and  $r = 1$  m, then

$$F = \frac{\mu_0}{2\pi} = 2 \times 10^{-7} \text{ N}$$

Thus, if two parallel wires carrying equal currents and placed apart in vacuum or air experience a mutual force of  $2 \times 10^{-7} \text{ N m}^{-1}$ , the current in each wire is said to be one ampere.

## 21.8) MOTION OF A CHARGED PARTICLE IN A UNIFORM FIELD

We can now think of various situations in which a moving charged particle or a current carrying conductor in a magnetic field experiences Lorentz force. The work done by a force on a body depends on its component in the direction of motion of the body. When the force on a charged particle in a magnetic field is perpendicular to its direction of motion, no work is said to be done. Hence the particle keeps the same speed and kinetic energy which it had while moving in the field, even though it is deflected. On the other hand, the speed and energy of a charged particle in an electrical field is always affected due to the force by the field on the particle. A charged particle moving perpendicular to a magnetic field follows a circular path as shown in Fig. 21.15 because it experiences a force at right angles to the direction of motion at every position.

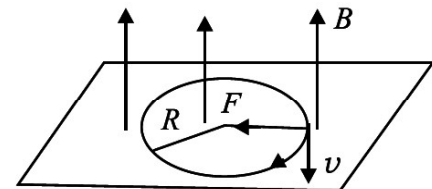


Fig. 21.15 : Path of a charged particle in a uniform magnetic field

To know the radius of the circular path of the charged particle, we note that the magnetic force ' $qvB$ ' provides the particle with the centripetal force  $\left(\frac{mv^2}{R}\right)$  that keeps it moving in a circle. So, we can write

$$qvB = \frac{mv^2}{R}$$

On rearranging, we get

$$R = \frac{mv}{qB} \quad (21.11)$$

The radius of the path traced by a charged particle in a uniform magnetic field is directly proportional to its momentum ( $mv$ ) and inversely proportional to its charge and the magnetic field. It means that greater the momentum, larger the circle, and stronger the field, the smaller the circle. The time period of rotation of the particle in a circular path is given by

$$T = \frac{2\pi R}{v} = \frac{2\pi m}{Bq} \quad (21.12)$$

Note that the time period is independent of velocity of the particle and radius of the orbit. It which means that once the particle is in the magnetic field, it would go round and round in a circle of the same radius. If  $m$ ,  $B$ ,  $q$  remain constant, the time period does not change even if  $v$  and  $R$  are changed.

Now think, what happens to  $R$  and  $T$  if (a) field  $B$  is made stronger, (b) field  $B$  is made weaker, (c) field  $B$  ceases to exist, (d) direction of  $B$  is changed, (e) the particle is made to enter the magnetic field at a higher speed, (f) the particle enters at an angle to  $B$  and (g) the charged particle loses its charge.

### Example 21.3

Refer to Fig. 21.16 and calculate the force between wires carrying current 10 A and 15 A, if their length is 5 m. What is the nature of this force?

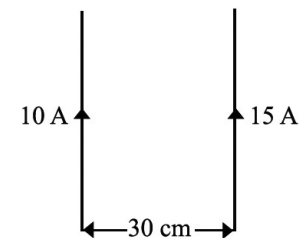


Fig. 21.16

### Solution :

When currents flow in two long parallel wires in the same direction, the wires attract each other and the force of attraction is given by

$$\frac{F}{5} = \frac{\mu_0 I_1 I_2}{2\pi r} = \frac{4\pi \times 10^{-7} \times 10 \times 15}{2\pi \times 0.3} = 10^{-4} \text{ Nm}^{-1}$$

$$\therefore F = 5 \times 10^{-4} \text{ Nm}^{-1}$$

The force is attractive in nature

### Example 21.4

An electron with velocity  $3 \times 10^7 \text{ ms}^{-1}$  describes a circular path in a uniform magnetic field of 0.2 T, perpendicular to it. Calculate the radius of the path.

### Solution :

We know that

$$R = \frac{mv}{qB}$$

Here,  $m_e = 9 \times 10^{-31}$  kg,  $e = 1.6 \times 10^{-19}$  C,  $v = 3 \times 10^7$  ms<sup>-1</sup> and  $B = 0.2$  T.

Hence,

$$R = \frac{9 \times 10^{-31} \times 3 \times 10^7}{0.2 \times 1.6 \times 10^{-19}} = 8.5 \times 10^{-4} \text{ m}$$

### Intext Questions 21.3

1. A stream of protons is moving parallel to a stream of electrons but in the opposite direction. What is the nature of force between them?
2. Both electrical and magnetic fields can deflect an electron. What is the difference between them?
3. A body is suspended from a vertical spring. What shall be the effect on the position of the body when a current is made to pass through the spring.

## 21.9) CURRENT LOOP AS A DIPOLE

From Eqn. 21.2, you will recall that the field at the centre of a coil is given by

$$B = \frac{\mu_0 I}{2r}$$

On multiplying the numerator and denominator by  $2\pi r^2$ , we can rewrite it as

$$B = \frac{\mu_0 I}{2r} \left( \frac{2\pi r^2}{2\pi r^2} \right)$$

$$B = \frac{\mu_0}{4\pi} \frac{2IA}{r^3}$$

$$B = \frac{\mu_0}{4\pi} \frac{2M}{r^3} \quad (21.13)$$

where  $A$  is area of coil and  $M (=AI)$  is magnetic moment. This shows that a current carrying coil behaves like a magnetic dipole having north and south poles. One face of the loop behaves as north pole while the other behaves as south pole.

**Let us now undertake a simple activity.**

### Activity 21.2

Suspend a bar magnet by a thread between pole pieces of a horse shoe magnet, as shown in Fig. 21.17.

What will happen when the bar magnet shown in Fig. 21.17 (a) is displaced slightly sideways? Since like poles repel, the bar magnet experiences a torque and tends to turn



through  $180^\circ$  and get aligned, as shown in Fig. 21.17 (b). Since a current loop behaves as a magnet, it will align in an external field in the same way.

## 21.10 THE ELECTROSTATIC ANALOG OF CURRENT LOOP

You have already studied the following equations in the lesson on electrostatics. The electric field of a dipole at a far point (say at  $r$  distance from its centre) on its axis is given by

$$E = \frac{1}{4\pi\epsilon_0} \frac{2P}{r^3} \quad (21.13 \text{ a})$$

The magnetic field due to a current carrying coil on its axis (at its centre and if  $r$  is the radius) is given by

$$B = \frac{\mu_0}{4\pi} \frac{2NIA}{r^3} = \frac{\mu_0}{4\pi} \frac{2M}{r^3} \quad (21.13 \text{ b})$$

where  $M$  is the magnetic dipole moment.

A comparison between these expressions leads us to the following analogies:

- A current loop behaves as a magnetic dipole with magnetic moment,
 
$$M = NIA \quad (21.13 \text{ c})$$
- Like the poles of a magnetic dipole, the two faces of a current loop are inseparable.
- A magnetic dipole in a uniform magnetic field behaves the same way as an electric dipole in a uniform electric field.
- A magnetic dipole also has a magnetic field around it similar to the electric field around an electric dipole.

Thus, magnetic field due to a magnetic dipole at an axial point of distance  $r$  from its centre is given by

$$B = \frac{\mu_0}{4\pi} \frac{2M}{r^3} \quad (21.13 \text{ d})$$

Whereas the field at an equatorial point is given by

$$B = -\frac{\mu_0}{4\pi} \frac{M}{r^3} \quad (21.13 \text{ e})$$

**The electrostatic analog:** If the above equation (Eqn. 21.13 d) derived for magnetic dipole is compared with the equation of electric dipole (Eqn. 21.13 a), we conclude that

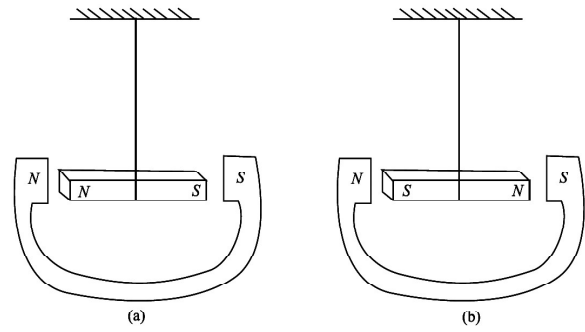


Fig. 21.17 : A bar magnet suspended between a horse shoe magnet.

the magnetic field due to a bar magnet of magnetic moment  $M$  at a large distance is analogous to that of an electric dipole of dipole moment  $P$  in an electric field. The corresponding relations can be stated as given below,

$$E \rightarrow B, \quad P \rightarrow M, \quad \frac{1}{4\pi\epsilon_0} \rightarrow \frac{\mu_0}{4\pi}$$

**Example 21.5**

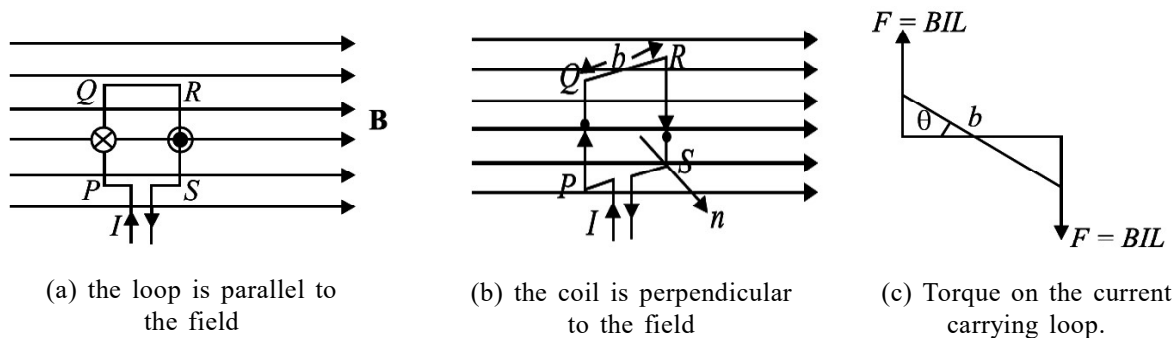
The smallest value of magnetic moment (associated with orbiting electron in atom) is called the Bohr Magneton  $\mu_B = \frac{eh}{4\pi m}$ . It is a fundamental constant. Calculate its value.

**Solution :**

$$\mu_B = \frac{eh}{4\pi m} = \frac{(1.6 \times 10^{-19} \text{ C}) \times (6.6 \times 10^{-34} \text{ Js})}{4 \times 3.14 \times (9 \times 10^{-31} \text{ kg})} = 9.34 \times 10^{-24} \text{ JT}^{-1}$$

**21.11 TORQUE ON A CURRENT LOOP**

A loop 'PQRS' ( $L$  &  $b$  are the length and breadth respectively) of current carrying wire placed in a uniform magnetic field ( $B$ ) experiences no net force but a torque act on it. This torque tends to rotate the loop to bring its plane perpendicular to the field direction. This is the principle that underlines the operation of all electric motors, meters etc. Let us examine the force on each side of a rectangular current carrying loop where plane is parallel to a uniform magnetic field  $B$ . Refer to Fig. 21.18 (a).



**Fig. 21.18 :** Force on the sides of a rectangular loop when

The sides 'PS' and 'QR' of the loop are parallel to  $B$ . So, no force will act on them. Sides 'PQ' and 'RS' are however, perpendicular to  $B$ , and these experience maximum force. We can easily find the direction of the force on 'PQ' and 'RS'.

In fact,  $|F_{PQ}| = |F_{RS}|$  and these act in opposite directions. Therefore, there is no net force on the loop. Since  $F_{PQ}$  and  $F_{RS}$  do not act along the same line, they exert a torque on the loop that tends to turn it. This holds good for a current loop of any shape in a magnetic field.

In case the plane of the loop is perpendicular to the magnetic field, there would neither be a net force nor a net torque on it [see Fig 21.18 (b)].

$$\begin{aligned} \text{Torque} &= \text{force} \times \text{perpendicular distance between the forces} \\ &= B I L b \sin \theta \end{aligned}$$

Refer to Fig. 21.18 (b) & (c) which shows a loop PQRS carrying current I.  $\theta$  is the angle between the magnetic field B and the normal ( $\vec{n}$ ) to the plane of the coil. Then, the torque is

$$\tau = N B I L b \sin \theta$$

where N is the number of turns of the coil. We can rewrite it as,

$$|\tau| = N B I A \sin \theta \quad (21.14)$$

where A (= L × b) is area of the coil.

$$|\tau| = B (NIA) \sin \theta$$

$$|\tau| = |B| |M| \sin \theta \quad (21.15)$$

Where M (= NIA) is known as the magnetic moment of the current carrying coil. Thus, we see that the torque depends on B, A, I, N and  $\theta$ .

If a uniform rotation of the loop is desired in a magnetic field, we need to have a constant torque. The couple would be approximately constant, if the plane of the coil were always along or parallel to the magnetic field. This is achieved by making the pole pieces of the magnet curved and placing a soft iron core at the centre so as to give a radial field. The soft iron core placed inside the loop would also make the magnetic field stronger and uniform resulting in greater torque (Refer to Fig. 21.19).

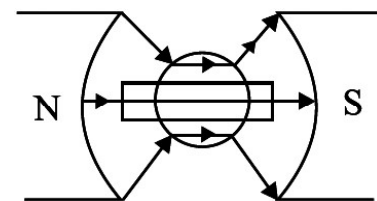


Fig. 21.19 : Constant torque on a coil in a radial field

## 21.12 GALVANOMETER

From what you have learnt so far, you can think of an instrument to detect current in any circuit. A device doing precisely this is called a galvanometer, which works on the principle that a current carrying coil, when placed in a magnetic field, experiences a torque.

A galvanometer consists of a coil wound on a non-magnetic frame. A soft iron cylinder is placed inside the coil. The assembly is supported on two pivots attached to springs with a pointer. This is placed between the pole pieces of a horse shoe magnet providing radial field (see Fig. 21.20).

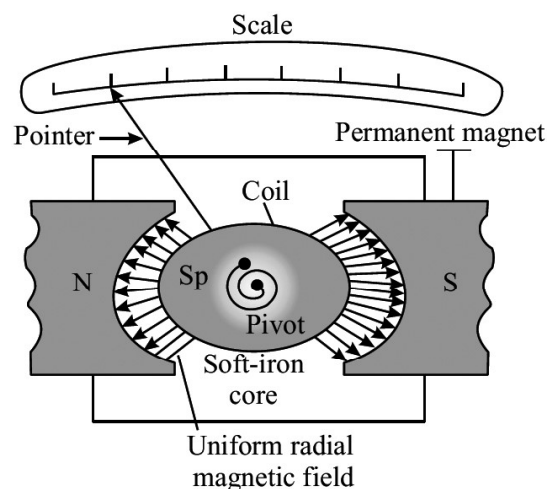


Fig. 21.20 : A moving coil galvanometer

To understand the working of a moving coil galvanometer, we recall that when a current is passed through the coil, it will rotate due to the torque acting on it. The spring sets up a restoring force and hence, a restoring torque. If  $\theta$  is the angle of twist and  $k$  is the restoring torque per unit twist or torsional constant, we can write  $NBIA \sin \theta = k\alpha$ . For  $\theta = 90^\circ$ , then  $\sin \theta = 1$ . So, in the instant case, we can write

$$NBIA = k\alpha$$

$$I = \frac{k\alpha}{NBA} \quad (21.16)$$

Where  $\frac{k}{NBA}$  is called galvanometer constant. From this we conclude that

$$\alpha \propto I$$

That is, deflection produced in a galvanometer is proportional to the current flowing through it provided  $N$ ,  $B$ ,  $A$  and  $k$  are constant. The ratio  $\frac{\alpha}{I}$  is known as current sensitivity of the galvanometer. It is defined as the deflection of the coil per unit current. The more the current stronger the torque and the coil turns more. Galvanometer can be constructed to respond to very small currents (of the order of  $0.1 \mu\text{A}$ ).

**Sensitivity of a galvanometer :** In order to have a more sensitive galvanometer,

- $N$  should be large
- $B$  should be large, uniform and radial
- area of the coil should be large and
- torsional constant should be small.

The values of  $N$  and  $A$  cannot be increased beyond a certain limit. Large values of  $N$  and  $A$  will increase the electrical and inertial resistance and the size of the galvanometer.  $B$  can be increased using a strong horse shoe magnet and by mounting the coil on a soft iron core. The value of  $k$  can be decreased by the use of materials such as quartz or phosphor bronze.

### 21.13 AN AMMETER AND A VOLTMETER

- (a) **Ammeter :** An Ammeter is a suitably shunted galvanometer. Its scale is calibrated to give the value of current in the circuit. To convert a galvanometer into an ammeter, a low resistance wire is connected in parallel with the galvanometer. The resistance of the shunt depends on the range of the ammeter and can be calculated as follows:

Let  $G$  be resistance of the galvanometer and  $N$  be the number of scale divisions in the galvanometer. Let  $k$  denote figure of merit or current for one scale deflection in the galvanometer. Then current which produces full scale deflection in the galvanometer is  $I_g = Nk$ .

Let  $I$  be the maximum current to be measured by the galvanometer.

Refer to Fig. 21.21. The voltage between points A and B is given by

$$V_{AB} = I_g G = (I - I_g)S$$

$$S = \frac{I_g G}{I - I_g} \quad (21.17)$$

Where  $S$  is the shunt resistance.

As  $G$  and  $S$  are in parallel, the effective resistance  $R$  of the ammeter is given by

$$R = \frac{GS}{G + S} \quad (21.18)$$

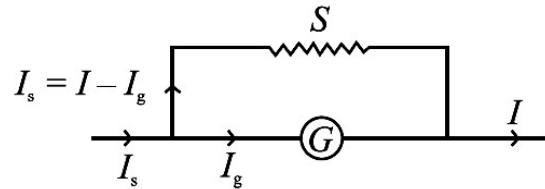


Fig. 21.21 : A shunted galvanometer acts as an ammeter

As the shunt resistance is small, the combined resistance of the galvanometer and the shunt is very low and hence, ammeter resistance is lower than that of the galvanometer. An ideal ammeter has almost negligible resistance. That is why when it is connected in series in a circuit, all the current passes through it without any observable drop.

- (b) **Voltmeter** : A voltmeter is used to measure the potential difference between two points in a circuit. We can convert a galvanometer into a voltmeter by connecting a high resistance in series with the galvanometer coil, as shown in Fig. 21.22. The value of the resistance depends on the range of voltmeter and can be calculated as follows:

A high resistance, say  $R$  is connected in series with the galvanometer coil. If the potential difference across  $AB$  is  $V$  volt, then total resistance of the voltmeter will be  $G + R$ . From Ohm's law, we can write

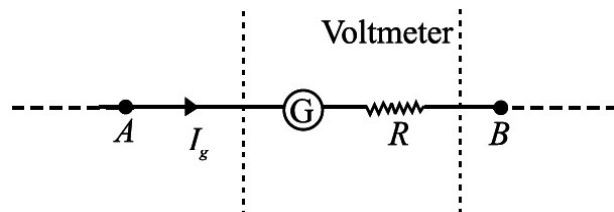


Fig. 21.22 : Galvanometer as a voltmeter

$$I_g (G + R) = V$$

$$G + R = \frac{V}{I_g}$$

$$R = \frac{V}{I_g} - G \quad (21.19)$$

This means that if a resistance  $R$  is connected in series with the coil of the galvanometer, it works as a voltmeter of range  $0 - V$  volts.

Now the same scale of the galvanometer which was recording the maximum potential ( $I_g \times G$ ) before conversion will record the potential  $V$  after conversion into voltmeter. The scale can be calibrated accordingly. The resistance of the voltmeter is higher than the resistance of galvanometer. Effective resistance of the voltmeter, is given by

$$R_v = R + G \quad (21.20)$$

The resistance of an ideal voltmeter is infinite. It is connected in parallel to the points across which potential drop is to be measured in a circuit. It will not draw any current. But the galvanometer coil deflects. Seems impossible! Think about it.

### Example 21.6

A circular coil of 30 turns and radius 8.0 cm, carrying a current of 6.0 A is suspended vertically in a uniform horizontal magnetic field of magnitude 1.0 T. The field lines make an angle of  $90^\circ$  with the normal to the coil. Calculate the magnitude of the counter torque that must be applied to prevent the coil from turning.

#### Solution :

Here,  $N = 30$ ,  $I = 6.0$  A,  $B = 1.0$  T,  $\theta = 90^\circ$ ,  $r = 8.0$  cm =  $8 \times 10^{-2}$  m

$$\text{Area (A) of the coil} = \pi r^2 = \frac{22}{7} \times (8 \times 10^{-2})^2 = 2.01 \times 10^{-2} \text{ m}^2$$

$$\begin{aligned} \text{Torque} &= N I A B \sin \theta \\ &= 30 \times 6 \times (2.01 \times 10^{-2}) \times 1 \times \sin 90^\circ \\ &= 30 \times 6 \times 2.01 \times 10^{-2} = 3.61 \text{ Nm} \end{aligned}$$

### Example 21.7

A galvanometer with a coil of resistance  $12.0 \Omega$  shows a full-scale deflection for a current of 2.5 mA. How will you convert it into (a) an ammeter of range 0 – 2A, and (b) voltmeter of range 0 – 10 volt?

#### Solution :

(a) Here,  $G = 12.0 \Omega$ ,  $I_g = 2.5$  mA =  $2.5 \times 10^{-3}$  A and  $I = 2$  A

From Eqn. 21.17, we have

$$\begin{aligned} S &= \frac{I_g G}{I - I_g} \\ &= \frac{2.5 \times 10^{-3} \times 12}{2 - (2.5 \times 10^{-3})} = 15 \times 10^{-3} \Omega \end{aligned}$$

So, for converting the galvanometer into an ammeter for reading 0 – 2A, a shunt of  $15 \times 10^{-3} \Omega$  resistance should be connected parallel to the coil.

(b) For conversion into voltmeter, let  $R$  be the resistance to be connected in series. From Eqn. 21.19, we have

$$\begin{aligned} R &= \frac{V}{I_g} - G \\ &= \frac{10}{2.5 \times 10^{-3}} - 12 = 4000 - 12 = 3988 \Omega \end{aligned}$$

Thus, a resistance of  $3988 \Omega$  should be connected in series to convert the galvanometer into voltmeter.

### Intext Questions 21.4

1. What is radial magnetic field?
2. What is the main function of a soft iron core in a moving coil galvanometer?
3. Which one has the lowest resistance - ammeter, voltmeter or galvanometer? Explain.
4. A galvanometer having a coil of resistance  $20 \Omega$  needs  $20 \text{ mA}$  current for full scale deflection. In order to pass a maximum current of  $3 \text{ A}$  through the galvanometer, what resistance should be added and how?

### WHAT YOU HAVE LEARNT

- Oersted experiment established that there is a magnetic field around a current carrying conductor.
- Every current carrying conductor develops a magnetic field around it. The magnetic field is given by Biot-Savart's Law,  $|\Delta B_0| = \frac{\mu_0}{4\pi} \frac{I \Delta l \sin \theta}{r^2}$ .
- Magnetic field at the centre of a coil carrying current is given by  $|B| = \frac{\mu_0 I}{2r}$ .
- Ampere's circuital law gives the magnitude of the magnetic field around a conductor,  $\oint B \cdot dl = \mu_0 I$
- Magnetic field due to an infinitely long current carrying conductor is  $|B| = \frac{\mu_0 I}{2\pi r}$ .
- Magnetic field inside a solenoid is  $|B| = \mu_0 nI$
- Magnetic field due to a solenoid at the ends is  $|B| = \frac{\mu_0 nI}{2}$ .

- The Lorentz force on a moving charge  $q$  is  $F = q(v \times B)$  and its direction is given by Fleming's left-hand rule.
- The mechanical force on a wire of length  $\Delta l$  and carrying a current of  $I$  in a magnetic field  $B$  is  $F = I \Delta l B$ .
- Mutual force per unit length between parallel straight conductors carrying currents  $I_1$  and  $I_2$  is given by  $\frac{F}{l} = \frac{\mu_0 I_1 I_2}{2\pi r}$ .
- If two parallel wires carrying equal currents and placed 1 m apart in vacuum or air experience a mutual force of  $2 \times 10^{-7} \text{ N m}^{-1}$ , the current in each wire is said to be one ampere.
- A charged particle moving perpendicular to a magnetic field traces a circular path of radius  $R = \frac{mv}{qB}$ .
- A current loop behaves like a magnetic dipole and its magnetic moment is  $M (= AI)$  where  $A$  is the loop area and  $I$  is the current flow through the loop.
- A current carrying coil placed in a magnetic field experiences a torque given by  $|\tau| = B (NIA) \sin \theta = |B| |M| \sin \theta$
- Galvanometer is used to detect electric current in a circuit.
- An ammeter is a shunted galvanometer and voltmeter is a galvanometer with a high resistance in series. Current is measured by an ammeter and potential difference by a voltmeter.

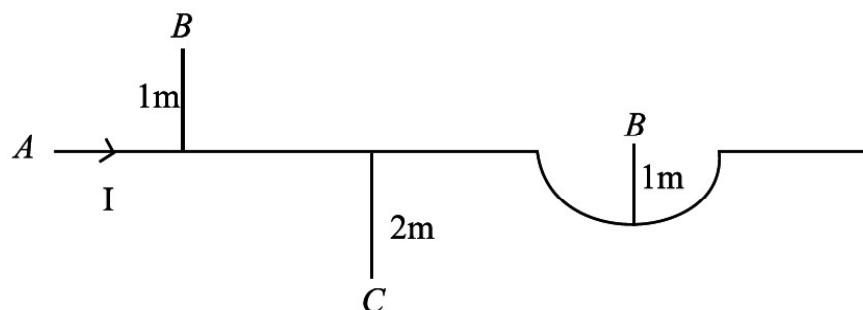
### TERMINAL EXERCISE

1. State and explain the Biot savart law.
2. Derive the expression for the magnetic field on the axis of a circular current loop.
3. State and explain the Ampere's circuital law
4. Derive the expression for Magnetic field due to an infinitely long current carrying conductor
5. Obtain the expression for magnetic field due to a solenoid
6. Obtain the expression for force acting on a charge moving in a magnetic field.
7. Derive the expression for the force per unit length acting between two parallel current carrying conductors and thus define the unit 'Ampere'.
8. Derive the expression for torque acting on a current loop in a uniform magnetic field.
9. Explain the principal, construction and working of a Galvanometer.



**TOSS**

10. How will you show that a current carrying conductor has a magnetic field around it? How will you find its magnitude and direction at a particular place?
11. A force acts upon a charged particle moving in a magnetic field, but this force does not change the speed of the particle, why?
12. At any instant a charged particle is moving parallel to a long, straight current carrying wire. Does it experience any force?
13. A current of 10 ampere is flowing through a wire. It is kept perpendicular to a magnetic field of 5T. Calculate the force on its  $\frac{1}{10}$ m length.
14. A long straight wire carries a current of 12 amperes. Calculate the intensity of the magnetic field at a distance of 48 cm from it.
15. Two parallel wires, each 3m long are situated at a distance of 0.05 m from each other. A current of 5 A flows in each of the wires in the same direction. Calculate the force acting on the wires. Comment on its nature?
16. The magnetic field at the centre of a 50 cm long solenoid is  $4.0 \times 10^{-2} \text{ NA}^{-1} \text{ m}^{-1}$  when a current of 8.0 A flows through it, calculate the number of turns in the solenoid.
17. Of the two identical galvanometers, one is to be converted into an ammeter and the other into a milliammeter. Which of the shunts will be of a larger resistance?
18. The resistance of a galvanometer is 20 ohms and gives a full-scale deflection for 0.005 A. Calculate the value of shunt required to change it into an ammeter to measure 1 A. What is the resistance of the ammeter?
19. An electron is moving in a circular orbit of radius  $5 \times 10^{-11} \text{ m}$  at the rate of  $7.0 \times 10^{15}$  revolutions per second. Calculate the magnetic fields B at the centre of the orbit.
20. Calculate the magnetic field at the centre of a flat circular coil containing 200 turns, of radius 0.16 m and carrying a current of 4.8 ampere.
21. Refer to Fig. 21.23 and calculate the magnetic field at A, B and C.



**Fig. 21.23**

## ANSWERS TO INTEXT QUESTIONS

### 21.1

1. (i) electrical (ii) magnetic as well as electrical.
2. A conductor in equilibrium is neutral i. e. it has no net electrical current. Due to their random motion, thermal electrons cancel the magnetic fields produced by them.
3. In first case length of wire  $l_1 = 2\pi r$ .

In second case length of wire  $l_2 = (2\pi r_2) \times 2 = 4\pi r_2$ .

$$\text{But } I_1 = I_2$$

$$\therefore 2\pi r = 4\pi r_2 \Rightarrow r_2 = \frac{r}{2}$$

$$\text{Using } |B| = \frac{\mu_0 n I}{2r}$$

$$|B_1| = \frac{\mu_0 I}{2r}, \quad |B_2| = \frac{\mu_0 \cdot 2 \cdot I}{2 \times \frac{r}{2}} = \frac{2\mu_0 I}{r} = 4B_1$$

That is, the magnetic field  $B$  at the centre of a coil with two turns is four times stronger than the field in first case.

### 21.2

1. C
2. Both laws specify magnetic field due to current carrying conductors.
3. (i) B, (ii) A, (iii) C.
4.  $B = \mu_0 \frac{n}{l} I \Rightarrow 4\pi \times \frac{10^{-7} \times n}{0.1\text{m}} \times 3\text{A} = 0.002$  or  $n = \frac{0.002 \times 10^7}{12\pi} = 50$  turns

### 21.3

1. The nature of the force will be attractive because the stream of protons is equivalent to electrons in the opposite direction.
2. The force exerted by a magnetic field on a moving charge is perpendicular to the motion of the charge and the work done by the force on the charge is zero. So, the KE of the charge does not change. In an electric field, the deflection is in the direction of the field. Hence the field accelerates it in the direction of field lines.

3. The direction of current in each turn of the spring is the same. Since parallel currents in the same direction exert force of attraction, the turns will come closer and the body shall be lifted upward, whatever be the direction of the current in the spring.

## 21.4

1. Radial magnetic field is one in which plane of the coil remains parallel to it.
2. This increases the strength of magnetic field due to the crowding of magnetic lines of force through the soft iron core, which in turn increases the sensitivity of the galvanometer.
3. Ammeter has the lowest resistance whereas voltmeter has the highest resistance. In an ammeter a low resistance is connected in parallel to the galvanometer coil whereas in a voltmeter, a high resistance is connected in series with it.
4. A low resistance  $R_s$  should be connected in parallel to the coil:

$$R_s = \frac{GI_g}{I - I_g} = \frac{20 \times (20 \times 10^{-3})}{3 - (20 \times 10^{-3})} = 0.13\Omega$$

## ANSWERS TO TERMINAL EXERCISE

13. 5 N
14. 5  $\mu\text{N}$
15. Attractive force of  $10^{-4} \text{ N m}^{-1}$
16.  $\frac{625}{\pi}$  turns
18. 0.1  $\Omega$
19. 4.48  $\pi\text{T}$
20. 1.2  $\pi \text{ mT}$
21.  $B_A = 2 \times 10^{-7} \text{ T}$ ,  $B_B = \pi \times 10^{-7} \text{ T}$  and  $B_C = 10^{-7} \text{ T}$



# ELECTROMAGNETIC INDUCTION

## INTRODUCTION

Electricity is the most convenient form of energy available to us. It lights our houses, runs trains, operates communication devices and makes our lives comfortable. The list of electrical appliances that we use in our homes is very long. Have you ever thought as to how is electricity produced?

Hydro-electricity is produced by a generator which is run by a turbine using the energy of water. In a coal, gas or nuclear fuel power station, the turbine uses steam to run the generator. Electricity reaches our homes through cables from the town substation. Have you ever visited an electric sub-station? What are the big machines installed there? These machines are called transformers. Generators and transformers are the devices, which basically make electricity easily available to us. These devices are based on the principle of electromagnetic induction.

In this lesson you will study electromagnetic induction, laws governing it and the devices based on it. You will also study the construction and working of electric generators, transformers and their role in providing electric power to us. A brief idea of eddy current and its application will also be undertaken in this chapter.

### OBJECTIVES

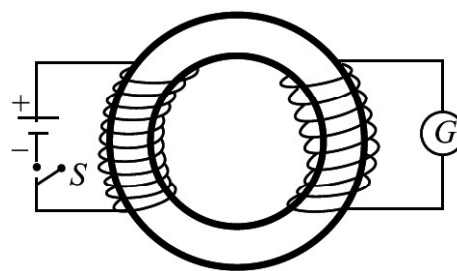
After studying this lesson, you should be able to

- *explain the phenomenon of electromagnetic induction with simple experiments;*
- *explain Faraday's and Lenz's laws of electromagnetic induction;*
- *know the concept of eddy currents;*
- *understand the concept of motional emf;*
- *describe the phenomena of self-induction and mutual induction;*
- *describe the working of ac and dc generators.*

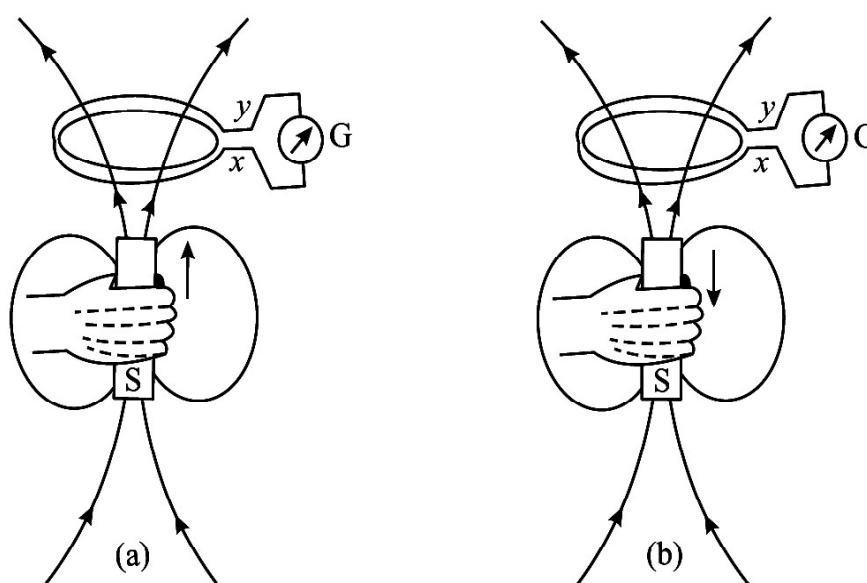
### 22.1) ELECTROMAGNETIC INDUCTION

In the previous lesson you have learnt that a steady current in a wire produces a steady magnetic field. Faraday initially (and mistakenly) thought that a steady magnetic field could produce electric current. Some of his investigations on magnetically induced currents used an arrangement similar to the one shown in Fig. 22.1. A current in the coil on the left

produces a magnetic field concentrated in the iron ring. The coil on the right is connected to a galvanometer  $G$ , which can indicate the presence of an induced current in that circuit. It is observed that there is no deflection in  $G$  for a steady current flow but when the switch  $S$  in the left circuit is closed, the galvanometer shows deflection for a moment. Similarly, when switch  $S$  is opened, momentary deflection is recorded but in opposite direction. It means that current is induced only when the magnetic field due to the current in the circuit on the left *changes*.



**Fig. 22.1** : Two coils are wrapped around an iron ring. The galvanometer  $G$  deflects for a moment when the switch is opened or closed



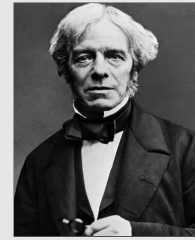
**Fig. 22.2** : Current induced in the coil when (a) magnet moves towards (b) away from the coil.

The importance of a change can also be demonstrated by the arrangement shown in Fig. 22.2. If the magnet is at rest relative to the coil, no current is induced in the coil. But when the magnet is moved towards the coil, current is induced in the direction indicated in Fig 22.2 (a). Similarly, if the magnet is moved away from the coil, the a current is induced in the opposite direction, as shown in Fig. 22.2 (b). Note that in both the cases, the magnetic field changes in the neighbourhood of the coil. An induced current is also observed to flow through the coil, if this is moved relative to the magnet. The presence of such currents in a circuit implies the existence of an *induced electromotive force (emf)* across the free ends of the coil, i.e.,  $x$  and  $y$ .

This phenomenon in which a magnetic field induces an emf is termed as *electromagnetic induction*. Faraday's genius recognised the significance of this work and followed it up. The quantitative description of this phenomenon is known as Faraday's law of electromagnetic induction. We will discuss it now.

### Michael Faraday (1791-1867)

British experimental scientist Michael Faraday is a classical example of a person who became great by sheer hardwork, perseverance, love for science and humanity. He started his carrier as an apprentice with a book binder, but utilized the opportunity to read science books that he received for binding. He sent his notes to Sir Humphry Davy, who immediately recognised the talent in the young man and appointed him his permanent assistant in the Royal Institute.



Sir Humphry Davy once admitted that the greatest discovery of his life was Michael Faraday. And he was right because Faraday made basic discoveries which led to the electrical age. It is because of his discoveries that electrical generators, transformers, electrical motors and electrolysis became possible.

#### 22.1.1 Faraday's Law of Electromagnetic Induction

The relationship between the changing magnetic field and the induced emf is expressed in terms of magnetic flux  $\phi_B$  linked with the surface of the coil. You will now ask: What is magnetic flux? To define **magnetic flux**  $\phi_B$  refer to Fig. 22.3 (a), which shows a typical infinitesimal element of area  $d\mathbf{s}$ , into which the given surface can be considered to be divided. The direction of  $d\mathbf{s}$  is normal to the surface at that point. By analogy with electrostatics, we can define the magnetic flux  $\phi_B$  for the area element  $d\mathbf{s}$  as

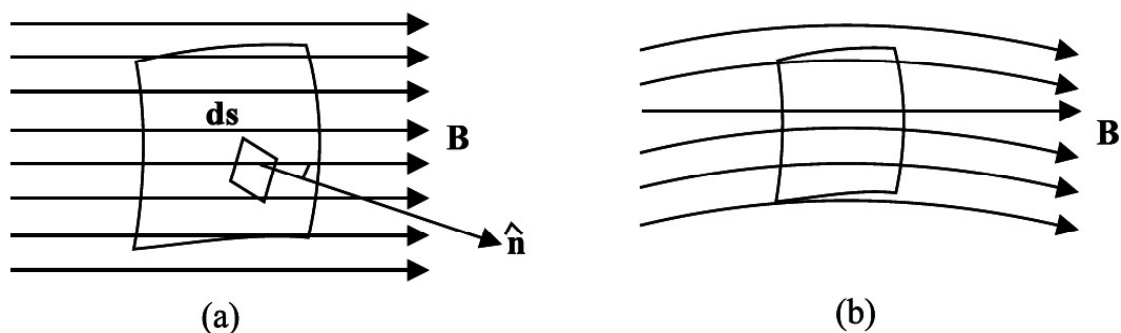
$$d\phi_B = \mathbf{B} \cdot d\mathbf{s} \quad (22.1)$$

The magnetic flux for the entire surface is obtained by summing such contributions over the surface. Thus,

$$d\phi_B = \Sigma \mathbf{B} \cdot d\mathbf{s} \quad (22.2)$$

The SI unit of magnetic flux is **weber** (Wb), where  $1 \text{ Wb} = 1 \text{ Tm}^2$ .

In analogy with electric lines and as shown in Fig. 22.3 (b), the number of magnetic lines intersecting a surface is proportional to the magnetic flux through the surface.



The magnetic flux for an infinitesimal area  $ds$  is given by  $d\phi_B = \mathbf{B} \cdot d\mathbf{s}$

The magnetic flux for a surface is proportional to the number of lines intersecting the surface

Fig. 22.3

**Faraday's law** states that an emf is induced across a loop of wire when the magnetic flux linked with the surface bound by the loop changes with time. The magnitude of induced emf is proportional to the rate of change of magnetic flux. Mathematically, we can write

$$|\varepsilon| = \frac{d\phi_B}{dt} \quad (22.3)$$

From this we note that weber (Wb), the unit of magnetic flux and volt (V), the unit of emf are related as  $1V = 1Wb \text{ s}^{-1}$ .

Now consider that an emf is induced in a closely wound coil. Each turn in such a coil behaves approximately as a single loop, and we can apply Faraday's law to determine the emf induced in each turn. Since the turns are in series, the total induced emf  $\varepsilon_r$  in a coil will be equal to the sum of the emfs induced in each turn. We suppose that the coil is so closely wound that the magnetic flux linking each turn of the coil has the same value at a given instant. Then the same emf  $\varepsilon$  is induced in each turn, and the total induced emf for a coil with N turns is given by

$$|\varepsilon_r| = N|\varepsilon| = N \left( \frac{d\phi_B}{dt} \right) \quad (22.4)$$

where  $\phi_B$  is the magnetic flux linked with a single turn of the coil.

Let us now apply Faraday's law to some concrete situations.

### Example 22.1

The axis of a 75 turn circular coil of radius 35 mm is parallel to a uniform magnetic field. The magnitude of the field changes at a constant rate from 25 mT to 50 mT in 250 milli second. Determine the magnitude of induced emf in the coil in this time interval.

#### Solution :

Since the magnetic field is uniform and parallel to the axis of the coil, the flux linking each turn is given by

$$\phi_B = B\pi R^2$$

where R is radius of a turn. Using Eqn. (22.4), we note that the induced emf in the coil is given by

$$|\varepsilon| = N \frac{d\phi_B}{dt} = N \frac{d(B\pi R^2)}{dt} = N\pi R^2 \frac{dB}{dt} = N\pi R^2 \left( \frac{B_2 - B_1}{t} \right)$$

Hence, the magnitude of the emf induced in the coil is

$$|\varepsilon| = 75\pi (0.035 \text{ m})^2 (0.10 \text{ Ts}^{-1}) = 0.30 \text{ V} = 30 \text{ mV}$$

This example explains the concept of emf induced by a time changing magnetic field.

**Example 22.2**

Consider a long solenoid with a cross-sectional area  $8 \text{ cm}^2$  [Fig. 22.4 (a) and 22.4 (b)]. A time dependent current in its windings creates a magnetic field  $B(t) = B_0 \sin 2\pi vt$ . Here  $B_0$  is constant, equal to  $1.2 \text{ T}$ . and  $v$ , the frequency of the magnetic field, is  $50 \text{ Hz}$ . If the ring resistance  $R = 1.0 \Omega$ , calculate the emf and the current induced in a ring of radius  $r$  concentric with the axis of the solenoid.

**Solution :**

We are told that magnetic flux

$$\phi_B = B_0 \sin 2\pi vt A$$

since normal to the cross sectional area of the solenoid is in the direction of magnetic field,

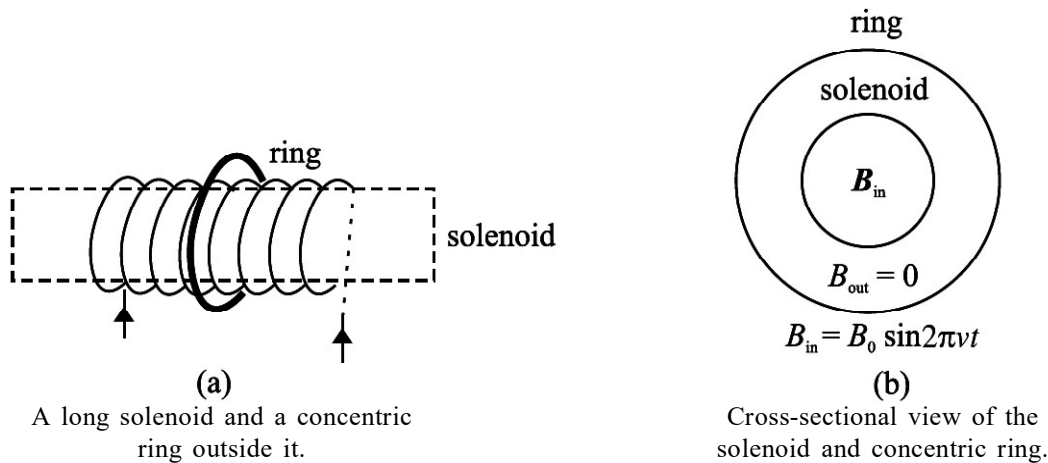


Fig. 22.4

$$\begin{aligned} \text{Hence, } |\varepsilon| &= \frac{d\phi_B}{dt} = 2\pi AB_0 \cos 2\pi vt \\ &= 2\pi \cdot (50 \text{ s}^{-1}) (8 \times 10^{-4} \text{ m}^2) (1.2 \text{ T}) \cos 2\pi vt \\ &= 0.3 \cos 2\pi vt \text{ volts} \\ &= 0.3 \cos 100\pi t \text{ volts} \end{aligned}$$

The current in the ring is  $I = \varepsilon/R$ . Therefore

$$\begin{aligned} I &= \frac{(0.3 \cos 100\pi t)}{(1.0\Omega)} \text{ volts} \\ &= + 0.3 \cos 100 \pi t \text{ amperes} \end{aligned}$$



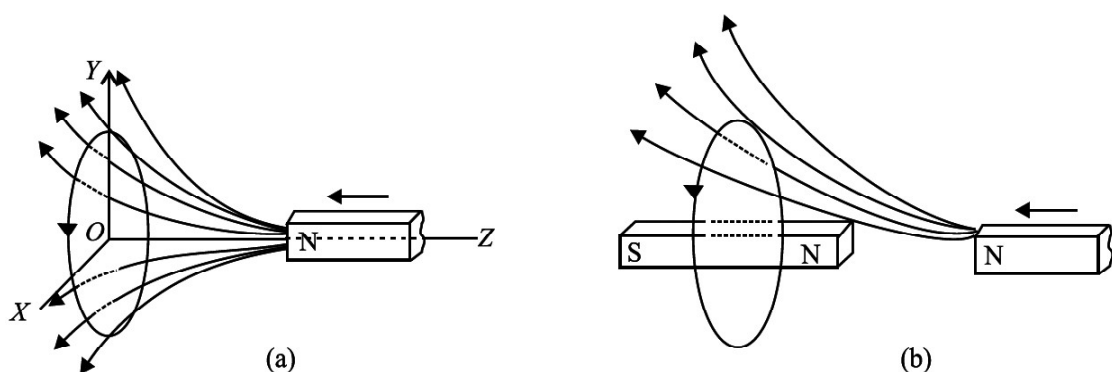
### Intext Questions 22.1

1. A 1000 turn coil has a radius of 5 cm. Calculate the emf developed across the coil if the magnetic field through the coil is reduced from 10 T to 0 in (a) 1 s (b) 1 ms.
2. The magnetic flux linking each loop of a 250-turn coil is given by  $\phi_b(t) = A + Dt^2$ , where  $A = 3 \text{ Wb}$  and  $D = 15 \text{ Wbs}^{-2}$  are constants. Show that (a) the magnitude of the induced emf in the coil is given by  $\varepsilon = (2ND)t$ , and (b) evaluate the emf induced in the coil at  $t = 0 \text{ s}$  and  $t = 3.0 \text{ s}$ .
3. The perpendicular to the plane of a conducting loop makes a fixed angle  $\theta$ , with a spatially uniform magnetic field. If the loop has area  $S$  and the magnitude of the field changes at a rate  $dB/dt$ , show that the magnitude of the induced emf in the loop is given by  $\varepsilon = (dB/dt) S \cos \theta$ . For what orientation(s) of the loop will  $\varepsilon$  be (a) maximum and (b) minimum?

#### 22.1.2 Lenz's Law

Consider a bar magnet approaching a conducting ring [Fig. 22.5 (a)]. To apply Faraday's law to this system, we first choose a positive direction with respect to the ring. Let us take the direction from  $O$  to  $Z$  as positive. (Any other choice is fine, as long as we are consistent.) For this configuration, the positive normal for the area of the ring is in the  $z$ -direction and the magnetic flux is negative. As the distance between the conducting ring and the N-pole of the bar magnet decreases, more and more field lines go through the ring, making the flux more and more negative. Thus  $d\phi_b/dt$  is negative. By Faraday's law,  $\varepsilon$  is positive relative to our chosen direction. The current  $I$  is directed as shown in Fig. 22.5.

The current induced in the ring creates a secondary magnetic field in it. This induced magnetic field can be taken as produced by a bar magnet, as shown in Fig. 22.5 (b). Recall that induced magnetic field repels or opposes the original magnetic field. This opposition is a consequence of the law of conservation of energy, and is formalized as Lenz's law. **When a current is induced in a conductor, the direction of the current will be such that its magnetic effect opposes the change that induced it.**



(a) A bar magnet approaching a metal ring

(b) the magnetic field of the induced current opposes the approaching bar magnet

Fig. 22.5

The key word in the statement is ‘oppose’-it tells us that we are not going to get something for nothing. When the bar magnet is pushed towards the ring, the current induced in the ring creates a magnetic field that opposes the change in flux. The magnetic field produced by the induced current repels the incoming magnet. If we wish to push the magnet towards the ring, we will have to do work on the magnet. This work shows up as electrical energy in the ring. Lenz’s law thus follows from the law of conservation of energy. We can express the combined form of Faraday’s and Lenz’s laws as

$$\varepsilon = -\frac{d\phi}{dt} \quad (22.5)$$

The negative sign signifies opposition to the cause.

As an application of Lenz’s law, let us reconsider the coil shown in example 22.2. Suppose that its axis is chosen in vertical direction and the magnetic field is directed along it in upward direction. To an observer located directly above the coil, what would be the sense of the induced emf? It will be clockwise because only then the magnetic field due to it (directed downward by the right-hand rule) will oppose the changing magnetic flux. You should learn to apply Lenz’s law before proceeding further. Try the following exercise.

Lenz’s law refers to induced currents in closed circuits. Because in closed circuits only the induced currents can be produced. Otherwise one can only speak of induced emf and its direction.

### Intext Questions 22.2

- The bar magnet in Fig. 22.6 moves to the right. What is the sense of the induced current in the stationary loop A? In loop B?

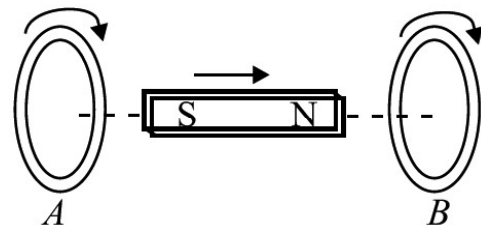


Fig. 22.6

- A cross-section of an ideal solenoid is shown in Fig. 22.7. The magnitude of a uniform magnetic field is increasing inside the solenoid and  $\mathbf{B} = 0$  outside the solenoid. In which conducting loops is there an induced current? What is the sense of the current in each case?

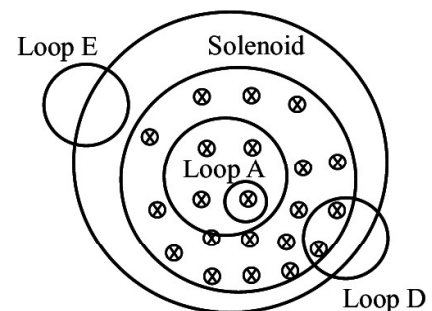


Fig. 22.7

- A bar magnet, with its axis aligned along the axis of a copper ring, is moved along its length toward the ring. Is there an induced current in the ring? Is there an induced electric field in the ring? Is there a magnetic force on the bar magnet? Explain.

## 22.2 MOTIONAL ELECTROMOTIVE FORCE

Consider a conductor, thick copper wire  $XY$  resting on a thick wire frame  $PQRS$  placed in a magnetic field  $B$ . Assume that there is no loss of energy due to friction. The magnetic field  $B$  is perpendicular to the plane of the closed area, acting downwards, as shown in the Fig. 22.8.  $XQRY$  forms a closed circuit enclosing an area that changes as  $XY$  moves.

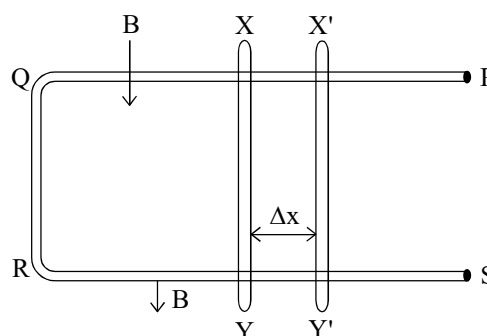


Fig. 22.8

Let the conductor of length  $l$  move to  $X'Y'$ , a distance  $\Delta x$  in time  $\Delta t$ . Then the increase in area of the loop is  $l\Delta x$ .

The change in magnetic flux  $\Delta\phi = B/l\Delta x$

$$\begin{aligned} \text{The rate of change in the magnetic flux} &= \frac{B/l\Delta x}{\Delta t} \\ &= B/lv \end{aligned}$$

where  $v$  is the velocity of the conductor  $\left( v = \frac{\Delta x}{\Delta t} \right)$

According to Faraday-Lenz's law, the induced emf is given by the rate of change of magnetic flux.

$$\therefore \quad \varepsilon = -\frac{\Delta\phi}{\Delta t} = -B/lv$$

Hence the numerical value of the induced emf generated in a conductor of length  $l$  moving with a velocity  $v$  placed in a magnetic field  $B$  is given by  $B/lv$ .

Thus we are able to produce induced emf by moving a conductor instead of varying the magnetic field.

This induced emf in a moving conductor in a magnetic field, is known as motional electromotive force.

### Intext Questions 22.3

1. A rod of length  $l$  moves with a constant speed  $v$  in a uniform perpendicular magnetic field  $B$ . Plot a graph showing the variation of induced emf with time.
2. A wire in N-S direction is dropped freely. What is the induced emf across ends of wire?

## 22.3 EDDY CURRENTS

When a solid mass of metal is moved in a magnetic field or placed near a varying magnetic flux, induced currents are set up throughout the interior of the metal in planes perpendicular to the lines of induction. These currents are known as eddy currents. The direction of these currents is governed by Lenz's law. This effect was discovered by Foucault in 1895.

The production of eddy currents can be shown by a simple experiment. A copper plate A [Fig. 22.9 (a)] which can swing about an axis perpendicular to its plane is made to oscillate between the poles of a powerful electromagnet. When the magnet is not energised (no current flows through electromagnet), the oscillations of the plate die down slowly under air damping. When the current is switched on through the electromagnet, the motion of the plate is stopped very quickly. This is because, the eddy currents set up in the plate circulate in a direction such as to oppose the motion of the plate (by Lenz's law).

If rectangular slots are made in the plate [Fig. 22.9 (b)], it oscillates for a longer time because the free circulation of Eddy currents then hindered by the air gaps.

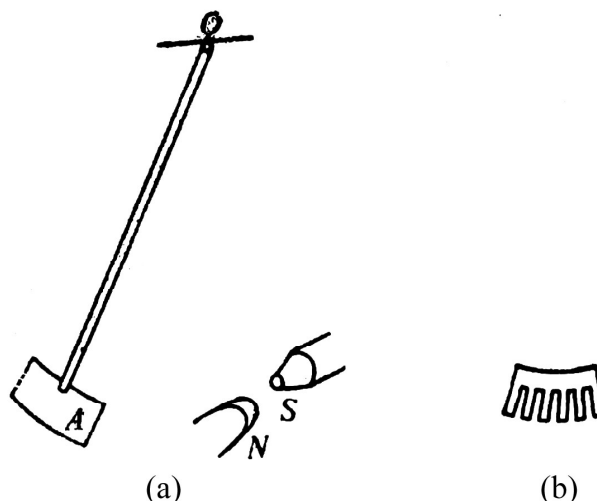


Fig. 22.9

Eddy currents are formed in the interior of the iron cores of the rotating armatures of motors and dynamos, and also in the cores of transformers. These currents are very undesirable both because of the large heating which they produce (loss of energy) and because of the flux which they themselves set up. Eddy currents are minimised by using laminations of metal to make a metal core. The laminations are separated by an insulating material like varnish. The plane of the laminations must be arranged parallel to the magnetic-field, so that they cut across the eddy current paths. This arrangement reduces the strength of the eddy currents.

Eddy currents are sometimes undesirable and sometimes usefully employed.

In a transformer, if we have a solid core the large eddy currents in the whole volume generate a lot of heat and thus waste a good amount of energy. Eddy currents also themselves generate magnetic flux, which is undesirable. Hence in transformers a laminated core, which increases the effective resistance and lowers the induced Eddy current is used; thus minimising the loss in energy.

### 22.3.1 Applications

- i) **Electric Brakes :** When a strong and stationary magnetic field is suddenly applied to a rotating drum, the eddy currents set up in the drum exert a torque which stops the motion of the drum. This principle is used in stopping electric trains.

- ii) **Induction Furnace** : In induction furnaces, the metal to be heated is placed in a rapidly changing magnetic field provided by a high-frequency alternating current. The eddy currents set up in the metal produce so much heat that the metal melts. This process is used in extracting the metal from ore.
- iii) **Induction Motor** : When a metallic cylinder is placed in a rotating magnetic field, eddy currents are set up in it. These currents try to reduce the relative motion between the cylinder and the field. Therefore, the cylinder begins to rotate in the direction of the field. This is the principle of an induction motor.

### Intext Questions 22.4

1. Name the instruments in which eddy currents concept is employed.
2. In the above instruments which type of metal plate is attached to the moving system?
3. In our domestic analog electric meters a rotating shiny disc is observed. Why?
4. List the undesirable effects of eddy currents.

## 22.4) INDUCTANCE

When current in a circuit changes, a changing magnetic field is produced around it. If a part of this field passes through the circuit itself, current is induced in it. Now suppose that another circuit is brought in the neighbourhood of this circuit. Then the magnetic field through that circuit also changes, inducing an emf across it. Thus, induced emfs can appear in these circuits in two ways:

- By changing current in a coil, the magnetic flux linked with each turn of the coil changes and hence an induced emf appears across that coil. This property is called **self-induction**.
- for a pair of coils situated close to each other such that the flux associated with one coil is linked through the other, a changing current in one coil induces an emf in the other. In this case, we speak of **mutual induction** of the pair of coils.



### 22.4.1 Self-Inductance

Let us consider a loop of a conducting material carrying electric current. The current produces a magnetic field **B**. The magnetic field gives rise to magnetic flux. The total magnetic flux linking the loop is

$$d\phi = \mathbf{B} \cdot d\mathbf{s}$$

In the absence of any external source of magnetic flux (for example, an adjacent coil carrying a current), the Biot-Savart's law tells us that the magnetic field and hence flux will be proportional to the current (I) in the loop, i.e.

$$\phi \propto I \quad \text{or} \quad \phi = LI \quad (22.6)$$

where  $L$  is called self-inductance of the coil. The circuit elements which oppose change in current are called **inductors**. These are in general, in the form of coils of varied shapes and sizes. The symbol for an **inductor** is . If the coil is wrapped around an iron core so as to enhance its magnetic effect, it is symbolised by putting two lines above it, as shown here . The inductance of an inductor depends on its geometry.

- (a) **Faraday's Law in terms of Self-Inductance:** So far you have learnt that if current in a loop changes, the magnetic flux linked through it also changes and gives rise to self-induced emf between the ends. In accordance with Lenz's law, the self-induced emf opposes the change that produces it. To express the combined form of Faraday's and Lenz's Laws of induction in terms of  $L$ , we combine Eqns. (22.5) and (22.6) to obtain

$$\varepsilon = -\frac{d\phi}{dt} = -L \frac{dI}{dt} \quad (22.7 \text{ a})$$

$$= -L \left( \frac{I_2 - I_1}{I} \right) \quad (22.7 \text{ b})$$

where  $I_1$  and  $I_2$  respectively denote the initial and final values of current at  $t = 0$  and  $t = \tau$ . Using Eqn. (22.7 b), we can define the unit of self-inductance:

$$\begin{aligned} \text{units of } L &= \frac{\text{unit of emf}}{\text{units of } dI/dt} \\ &= \frac{\text{volt}}{\text{ampere/second}} \\ &= \text{ohm-second} \end{aligned}$$

An ohm-second is called a *henry*, (abbreviated H). For most applications, henry is a rather large unit, and we often use millihenry, mH ( $10^{-3}$  H) and microhenry  $\mu\text{H}$  ( $10^{-6}$ H) as more convenient measures.

The self-induced emf is also called the **back emf**. Eqn. (22.7 a) tells us that the **back emf in an inductor** depends on the rate of change of current in it and **opposes the change in current**. Moreover, since an infinite emf is not possible, from Eqn. (22.7 b) we can say that an instantaneous change in the inductor current cannot occur. Thus, we conclude that **current through an inductor cannot change** instantaneously.

The inductance of an inductor depends on its geometry. In principle, we can calculate the self-inductance of any circuit, but in practice it is difficult except for devices with simple geometry. A solenoid is one such device used widely in electrical circuits as inductor. Let us calculate the self-inductance of a solenoid.

- (b) **Self-inductance of a solenoid :** Consider a long solenoid of cross-sectional area  $A$  and length  $l$ , which consists of  $N$  turns of wire. To find its inductance, we must

relate the current in the solenoid to the magnetic flux through it. In the preceding lesson, you used Ampere's law to determine magnetic field of a long solenoid:

$$|\mathbf{B}| = \mu_0 n I$$

where  $n = N/l$  denotes is the number of turns per unit length and  $I$  is the current through the solenoid.

The total flux through  $N$  turns of the solenoid is

$$\phi = N |\mathbf{B}| A = \frac{\mu_0 N^2 A I}{l} \quad (22.8)$$

and self-inductance of the solenoid is

$$L = \frac{\phi}{I} = \frac{\mu_0 N^2 A}{l} \quad (22.9)$$

Using this expression, you can calculate self-inductance and back emf for a typical solenoid to get an idea of their magnitudes.

### Intext Questions 22.5

1. A solenoid 1 m long and 20 cm in diameter contains 10,000 turns of wire. A current of 2.5 A flowing in it is reduced steadily to zero in 1.0 ms. Calculate the magnitude of back emf of the inductor while the current is being reduced.
2. A certain length ( $l$ ) of wire, folded into two parallel, adjacent strands of length  $l/2$ , is wound on to a cylindrical insulator to form a type of wire-wound non-inductive resistor (Fig. 22.10). Why is this configuration called non-inductive?
3. What rate of change of current in a 9.7 mH solenoid will produce a self-induced emf of 35 mV?

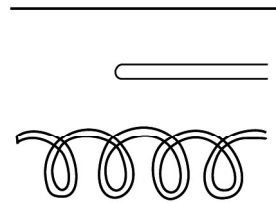


Fig. 22.10

### 22.4.2 Mutual Inductance

When current changes in a coil, a changing magnetic flux develops around it, which may induce emf across an adjoining coil. As we see in Fig. 22.11, the magnetic flux linking each turn of coil  $B$  is due to the magnetic field of the current in coil  $A$ .

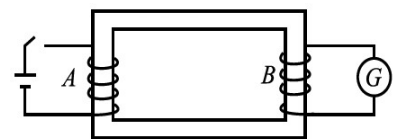


Fig. 22.11 : Mutual inductance of a pair of coils

Therefore, a changing current in each coil induces an emf in the other coil,

$$\text{i.e.,} \quad \phi_2 \propto \phi_1 \propto I_1 \Rightarrow \phi_2 = M I_1 \quad (22.10)$$

where  $M$  is called the mutual inductance of the pair of coils. Also back emf induced across the second coil

$$\varepsilon_2 = - \frac{d\phi}{dt}$$

$$= -M \frac{dI}{dt} = -M \left( \frac{I_2 - I_1}{t} \right) \quad (22.11)$$

where the current in coil A changes from  $I_1$  to  $I_2$  in  $t$  seconds.

The mutual inductance depends only on the geometry of the two coils, if no magnetic materials are nearby. The SI unit of mutual inductance is also henry (H), the same as the unit of self-inductance.

### Example 22.3

A coil in one circuit is close to another coil in a separate circuit. The mutual inductance of the combination is 340 mH. During a 15 ms time interval, the current in coil 1 changes steadily from 28 mA to 57 mA and the current in coil 2 changes steadily from 36 mA to 16 mA. Determine the emf induced in each coil by the changing current in the other coil.

#### Solution :

During the 15 ms time interval, the currents in the coils change at the constant rates of

$$\frac{di_1}{dt} = \frac{57 \text{ mA} - 23 \text{ mA}}{15 \text{ ms}} = 2.3 \text{ As}^{-1}$$

$$\frac{di_2}{dt} = \frac{16 \text{ mA} - 36 \text{ mA}}{15 \text{ ms}} = -1.3 \text{ As}^{-1}$$

From Eqn. (22.11), we note that the magnitudes of the induced emfs are

$$\varepsilon_1 = - (340 \text{ mH}) (2.3 \text{ As}^{-1}) = - 0.78 \text{ V}$$

$$\varepsilon_2 = (340 \text{ mH}) (1.3 \text{ As}^{-1}) = 0.44 \text{ V}$$

Remember that the minus signs in Eqn. (22.11) refer to the sense of each induced emf.

One of the most important applications based on the phenomenon of mutual inductance is transformer. You will learn about it later in this lesson. Some commonly used devices based on self-inductance are the choke coil and the ignition coil. We will discuss about these devices briefly. Later, you will also learn that a combination of inductor and capacitor acts as a basic oscillator. Once the capacitor is charged, the charge in this arrangement oscillates between its two plates through the inductor.

### Intext Questions 22.6

1. Consider the sense of the mutually induced emf's in Fig. 22.11, according to an observer located to the right of the coils. (a) At an instant when the current  $i_1$  is increasing, what is the sense of emf across the second coil? (b) At an instant when  $i_2$  is decreasing, what is the sense of emf across the first coil?
2. Suppose that one of the coils in Fig. 22.11 is rotated so that the axes of the coils are perpendicular to each other. Would the mutual inductance remain the same, increase or decrease? Explain.



## 22.5) POWER GENERATOR

One of the most important sources of electrical power is called **generator**. A generator is a device that converts mechanical energy into electrical energy with the help of magnetic field. No other source of electric power can produce as large amounts of electric power as the generator. A conductor or a set of conductors is rotated in a magnetic field and voltage is developed across the rotating conductor due to electromagnetic induction. The energy for the rotation of the conductors can be supplied by water, coal, diesel or gas or even nuclear fuel. Accordingly, we have hydro-generators, thermal generators, and nuclear reactors, respectively.

There are two types of generators

- alternating current generator or A.C. generator also called alternators.
- direct current generator or D.C. generator or dynamo.

Both these generators work on the principle of electromagnetic induction.

### 22.5.1 A.C. Generator or Alternator

A generator basically consists of a loop of wire rotating in a magnetic field. Refer to Fig. 22.12. It shows a rectangular loop of wire placed in a uniform magnetic field. As the loop is rotated along a horizontal axis, the magnetic flux through the loop changes. To see this, recall that the magnetic flux through the loop, as shown in Fig. 22.12, is given by

$$\phi(t) = \mathbf{B} \cdot \hat{\mathbf{n}} A$$

where  $\mathbf{B}$  is the field,  $\hat{\mathbf{n}}$  is a unit vector normal to the plane of the loop of area  $A$ . If the angle between the field direction and the loop at any instant is denoted by  $\theta$ ,  $\phi(t)$  can be written as

$$\phi(t) = AB \cos \theta$$

When we rotate the loop with a constant angular velocity  $\omega$ , the angle  $\theta$  changes as

$$\theta = \omega t \quad (22.12)$$

∴

$$\phi(t) = AB \cos \omega t$$

Now, using Faraday's law of electromagnetic induction, we can calculate the emf induced in the loop :

$$\varepsilon(t) = -\frac{d\phi}{dt} = \omega AB \sin \omega t \quad (22.13)$$

The emf induced across a coil with  $N$  number of turns is given by

$$\begin{aligned} \varepsilon(t) &= N \omega AB \sin \omega t \\ &= \varepsilon_0 \sin \omega t \end{aligned} \quad (22.14)$$

where  $\varepsilon_0$  is maximum emf

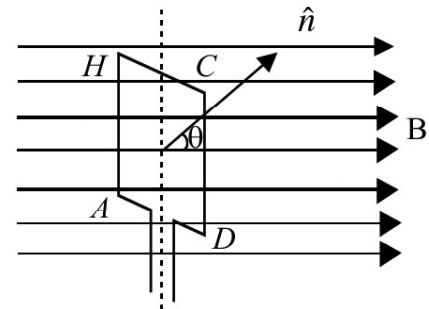


Fig. 22.12 : A loop of wire rotating in a magnetic field.

That is, when a rectangular coil rotates in a uniform magnetic field, the induced emf is sinusoidal.

An A.C. generator consists of four main parts (see in Fig. 22.13) : (i) Armature, (ii) Field magnet, (iii) Slip-rings, (iv) Brushes.

An armature is a coil of large number of turns of insulated copper wire wound on a cylindrical soft iron drum. It is capable of rotation at right angles to the magnetic field on a rotor shaft passing through it along the axis of the drum. This drum of soft iron serves two purpose : it supports the coil, and increases magnetic induction through the coil. A field magnet provided to produce a uniform and permanent radial magnetic field between its pole pieces.

Slip Rings provide alternating current generated in armature to flow in the device connected across them through brushes. These are two metal rings to which the two ends of the armatures are connected. These rings are fixed to the shaft. They are insulated from the shaft as well as from each other. Brushes are two flexible metal or carbon rods [ $B_1$  and  $B_2$  (Fig. 22.13)], which are fixed and constantly in touch with revolving rings. It is with the help of these brushes that the current is passed on from the armature and rings to the main wires which supply the current to the outer circuit.

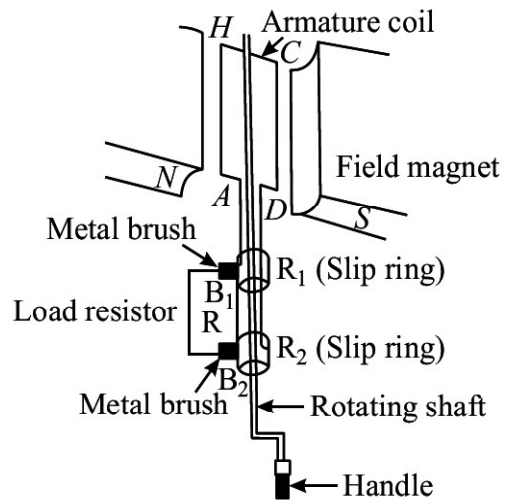


Fig. 22.13 : Schematics of an ac generator

The principle of working of an ac generator is illustrated in Fig. 22.14.

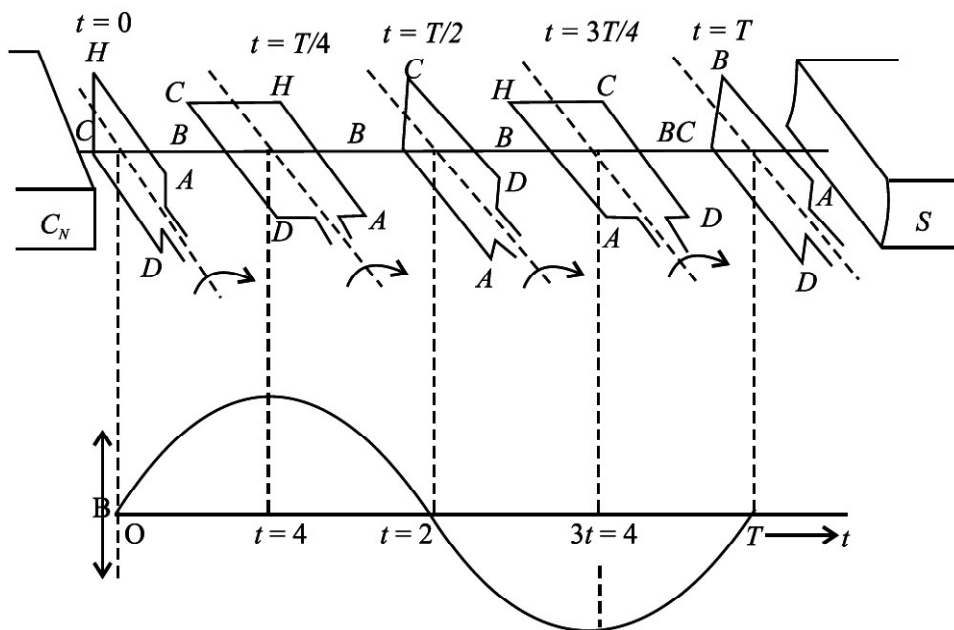


Fig. 22.14 : Working principle of an ac generator

Suppose the armature coil AHCD rotates in the anticlockwise direction. As it rotates, the magnetic flux linked with it changes and the current is induced in the coil. The direction of the induced current is given by Fleming's right hand rule. Considering the armature to be in the vertical position and its rotation in anticlockwise direction, the wire AH moves downward and DC moves upwards, the direction of induced emf is from H to A and D to C i.e., in the coil it flows along DCHA. In the external circuit the current flows along  $B_1 R B_2$  as shown in Fig. 22.14 (a). This direction of current remains the same during the first half turn of the armature. However, during the second half revolution (Fig. 22.14 b), the wire AH moves upwards while the wires CD moves downwards. The current flows in the direction AHCD in the armature coil i.e., the direction of induced current in the coil is reversed. In the external circuit direction is  $B_2 R B_1$ . Therefore, the direction of the induced emf and the current changes after every half revolution in the external circuit also. Hence, the current thus produced alternates in each cycle (Fig. 22.14 c).

The arrangement of slip rings and brushes creates problems of insulation and sparking when large output powers are involved. Therefore, in most practical generators, the field is rotated and the armature (coil) is kept stationary. In such a generator, armature coils are fixed permanently around the inner circumference of the housing of the generator while the field coil pole pieces are rotated on a shaft within the stationary armature.

### 22.5.2 Dynamo (DC Generator)

*A dynamo is a machine in which mechanical energy is changed into electrical energy in the form of direct current.* You must have seen a dynamo attached to a bicycle for lighting purpose. In automobiles, dynamo has a dual function for lighting and charging the battery. The essential parts of dynamo are (i) field magnet, (ii) armature, (iii) commutator split rings and (iv) brushes.

*Armatures and field magnets differ in dynamo and alternator. In the dynamo, the field magnets are stationary and the armature rotates while in an alternator, armature is stationary (stator) and the field magnet (rotor) rotates.*

In a dynamo, ac waveform or the sine wave produced by an a.c. generator is converted into d.c. form by the split ring commutator. Each half of the commutator is connected permanently to one end of the loop and the commutator rotates with the loop. Each brush presses against one segment of the commutator. The brushes remain stationary while the commutator rotates. The brushes press against opposite segments of the commutator and every time the voltage reverses polarity, the split rings change position. This means that one brush always remains positive while the other becomes negative, and a d.c. fluctuating voltage is obtained across the brushes.

A dynamo has almost the same parts as an ac dynamo but it differs from the latter in one respect: In place of slip ring, we put two split rings  $R_1$  and  $R_2$  which are the two half of the same ring, as shown in Fig. 22.15 (a). The ends of the armature coil are connected to these rings and the ring rotates with the armature and changes the contact with the brushes  $B_1$  and  $B_2$ . This part of the dynamo is known as **commutator**.

When the coil is rotated in the clockwise direction, the current produced in the armature is a.c. but the commutator changes it into d.c. in the outer circuit. In the first half cycle,

Fig. 22.15 (a), current flows along DCHA. The current in the external circuit flows along  $B_1$  L  $B_2$ . In the second half, Fig. 22.15 (b), current in the armature is reversed and flows along AHCD and as the ring  $R_1$  comes into contact with  $B_1$  to  $B_2$ . Thus, current in the external circuit always flows in the same direction. The current produced in the outer circuit is graphically represented in Fig. 22.15 (c) as the coil is rotated from the vertical position, perpendicular to the magnetic lines of force. The current generated by such a simple d.c. dynamo is unidirectional but its value varies considerably and even falls to zero twice during each rotation of the coil.

One way of overcoming this variation would be to use two coils, mutually at right angles, and to divide the commutator ring into four sections, connected to the ends of the coils. In such a case, both these coils produce emf of the same type but they differ in phase by  $\pi/2$ . The resultant current or emf is obtained by superposition of the two, as shown in Fig. 22.15 (d). In this way, the fluctuations are considerably reduced. Similarly, in order to get a steady current, we use a large number of coils, each consisting of good many turns. The commutator ring is divided into as many segments as the number of ends of coils, so that the coils work independently and send current into the outer circuit. The resultant current obtained is shown in Fig. 22.15 (e) which is practically parallel to the time axes.

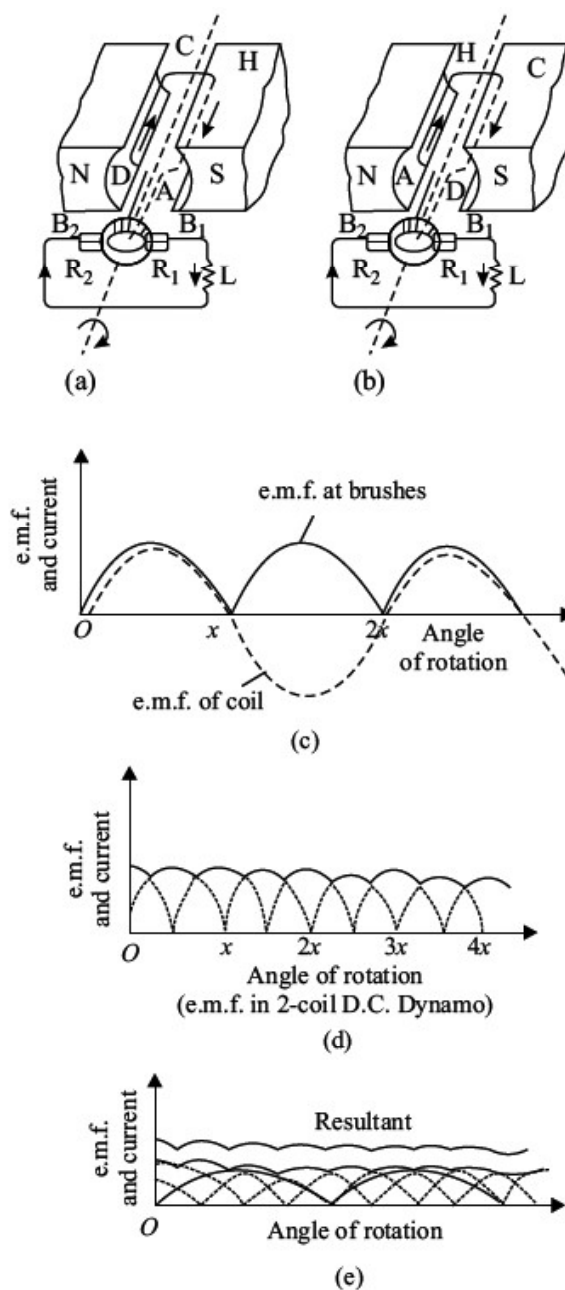


Fig. 22.15 : A dc generator

### Intext Questions 22.7

1. Distinguish between an ac and dc generator.
2. Name the essential parts of a generator.
3. Why do we use a commutator in a dc generator.
4. Where do you find the use of dynamo in daily life?

## WHAT YOU HAVE LEARNT

- A current is induced in a coil of wire if magnetic flux linking the surface of the coil changes. This is known as the phenomenon of *electromagnetic* induction.
- The induced emf  $\varepsilon$  in a single loop is given by *Faraday's law*:

$$\varepsilon = \frac{d\phi_B}{dt}$$

where  $\phi_B$  is the magnetic flux linking the loop.

- According to *Lenz's Law*, the induced emf opposes the cause which produces it.
- If the current changes in a coil, a self-induced emf exists across it.
- For a long, tightly wound solenoid of length  $l$ , cross-sectional area  $A$ , having  $N$  number of turns, the self-inductance is given by

$$L = \frac{\mu_0 N^2 A}{l}$$

- The changing currents in two nearby coils induce emf mutually.
- Whenever there is a change in the magnetic flux linked with a circuit or a metal block, there is an induced current produced in the circuit or block. This is known as eddy current.
- Whenever a conductor is moved with a uniform speed in the region of a uniform magnetic field, an induced emf is produced across the ends. This induced emf is known as motional emf.
- A generator converts mechanical energy into electrical energy. It works on the principle of electromagnetic induction.

## TERMINAL EXERCISE

1. Each loop in a 250-turn coil has face area  $S = 9.0 \times 10^{-2} \text{ m}^2$ . (a) What is the rate of change of the flux linking each turn of the coil if the induced emf in the coil is 7.5 V? (b) If the flux is due to a uniform magnetic field at  $45^\circ$  from the axis of the coil, calculate the rate of change of the field to induce that emf.
2. (a) In Fig. 22.16 (a) what is the direction of the induced current in the loop when the area of the loop is decreased by pulling on it with the forces labelled F? B is directed into the page and perpendicular to it.  
(b) What is the direction of the induced current in the smaller loop of Fig. 22.16 (b) when a clockwise current as seen from the left is suddenly established in the larger loop, by a battery not shown?

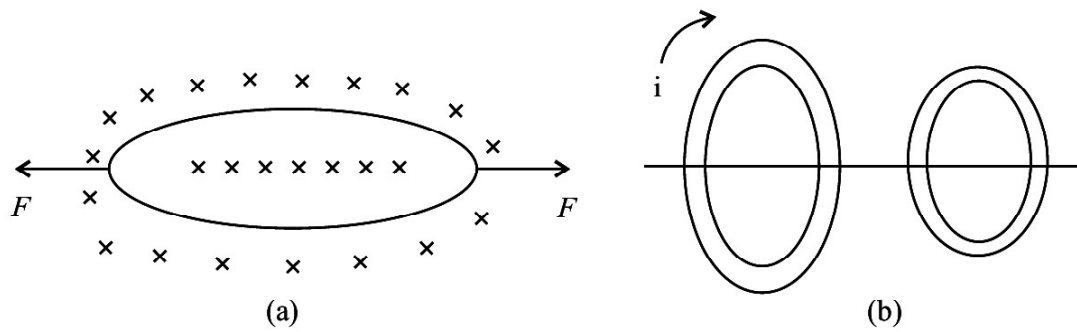


Fig. 22.16

3. (a) If the number of turns in a solenoid is doubled, by what amount will its self-inductance change?
- (b) Petrol in a vehicle's engine is ignited when a high voltage applied to a spark plug causes a spark to jump between two conductors of the plug. This high voltage is provided by an ignition coil, which is an arrangement of two coils wound tightly one on top of the other. Current from the vehicle's battery flows through the coil with fewer turns. This current is interrupted periodically by a switch. The sudden change in current induces a large emf in the coil with more turns, and this emf drives the spark. A typical ignition coil draws a current of 3.0 A and supplies an emf of 24 kV to the spark plugs. If the current in the coil is interrupted every 0.10 ms, what is the mutual inductance of the ignition coil?
4. What is electromagnetic induction? Explain Faraday's laws of electromagnetic induction.
5. State Lenz's law. Show that Lenz's law is a consequence of law of conservation of energy.
6. What is self-induction? Explain the physical significance of self-inductance.
7. Distinguish between the self-inductance and mutual-inductance. On what factors do they depend?
8. How much e.m.f. will be induced in a 10 H inductor in which the current changes from 10 A to 7 A in  $9 \times 10^{-2}$  s?
9. Deduce an expression for motional emf induced in a rod when it is placed over a thick frame in a uniform magnetic field.
10. A rod of length 80 cm moves with a constant speed of 4 m/s in a region having a constant perpendicular magnetic field. The induced emf across the ends of rod is 64 mV. What is B value?
11. What are eddy currents? Explain. What are important applications of eddy currents?
12. Give the principle, construction and explain working of an a.c. generator.
13. Give the principle, construction and explain working of a d.c. generator.
14. How many volts are generated in a wire 10 cm long which cuts directly across a magnetic field of flux density 2 web/m<sup>2</sup>, if it moves at a speed of 4 m/sec?

## ANSWERS TO INTEXT QUESTIONS

### 22.1

1.  $N = 1000$ ,  $r = 5 \times 10^{-2}$  m and  $B_1 = 10$  T  $B_2 = 0$ T

(a) For  $t = 1$  s

$$\begin{aligned} |\varepsilon| &= N \frac{(B_2 - B_1)}{t} \pi r^2 \\ &= 10^3 \times \frac{10 \times \pi \times 25 \times 10^{-4}}{1} \\ &= 25 \pi \text{V} \\ &= 25 \times 3.14 = 78.50 \text{ V} \end{aligned}$$

(b) For  $t = 1$  ms

$$\begin{aligned} |\varepsilon| &= \frac{10^3 \times 10\pi \times 25 \times 10^{-4}}{10^{-3}} \\ &= 78.5 \times 10^3 \text{V} \end{aligned}$$

2. Since  $\phi = A + Dt^2$ ,  $\varepsilon_1 = \frac{d\phi}{dt} = 2Dt$

$$\begin{aligned} \therefore \varepsilon &= N\varepsilon_1 = 2N Dt \\ &= 2 \times 250 \times 15 t = 7500 t \end{aligned}$$

For  $t = 0$ ,  $\varepsilon_1 = 0$  and hence  $\varepsilon = 0$  V

For  $t = 3$  s,  $\varepsilon = 22500$  V

3.  $\phi = \mathbf{B} \cdot \mathbf{S} = BS \cos \theta$

$$|\varepsilon| = N \frac{d\phi}{dt}$$

$$|\varepsilon| = \left| NS \frac{dB}{dt} \cos \theta \right| \because \theta \text{ is const}$$

(a)  $|\varepsilon|$  is max.

when  $\cos \theta = 1$ ,  $\theta = 0$ , i.e., The coil is normal to the field.

(b)  $|\varepsilon|$  is min.

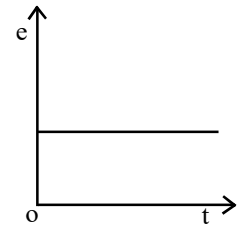
when  $\theta = 90$ , i.e., coil surface is parallel to the field.

**22.2**

1. As we look on the coil from magnet side, anticlockwise for both A and B.
2. In all the loops except loop E there is a change in magnetic flux. For each of them the induced current will be anticlockwise.
3. Yes, there is an induced current in the ring. The bar magnet is acted upon by a repulsive force due to the induced current in the ring.

**22.3**

1. Since the rod moves with a constant speed, the induced emf ( $\epsilon$ ) is constant and equals to  $Blv$ .
2. As there is no magnetic field in a direction perpendicular to the direction of motion of wire, induced emf across the ends of wire is zero.


**22.4**

1. Moving coil Galvanometer, balances, ammeters and Voltmeters.
2. Copper plate
3. Electric currents are induced in the disc by magnetic fields produced by sinusoidally varying currents in a coil.
4. (i) Eddy currents opposes relative motion.  
(ii) Eddy current involves loss of energy in the form of heat. Consequently, the life of appliance is reduced.

**22.5**

$$\begin{aligned}
 1. \quad \epsilon &= L \frac{dI}{dt} = \omega \frac{N^2 A}{l} \frac{(I_2 - I_1)}{t} \\
 &= \frac{4\pi \times 10^{-7} \times \pi \times 10^{-2} \times (2.5 - 0)}{1 \times 10^{-3}} \\
 &= 10^{-6} \text{ V}
 \end{aligned}$$

2. Because, current in the two parallel strands flow in opposite directions and oppose the self induced currents and thus minimize the induction effects.



$$3. \quad 3.5 \times 10^{-3} = 9.7 \times 10^{-3} \times \frac{dI}{dt}$$

$$= \frac{dI}{dt} = \frac{3.5}{9.7} = 0.36 \text{ A s}^{-1}$$

**22.6**

1. (a) If  $i_1$  is increasing, the flux emerging out of the first coil is also increasing. Therefore, the induced current in the second coil will oppose this flux by a current flowing in clockwise sense as seen by O. Therefore B will be positive and A negative.
- (b) If  $i_2$  is decreasing, flux emerging out of the first coil is decreasing. To increase it the induced current should flow in out anticlockwise sense leaving C at positive potential and D at negative.
2. No, the mutual inductance will decrease. Because, when the two coils are at right angles, coupling of flux from one coil to another coil will be the least.

**22.7**

1. (i) The a.c. generator has slip rings whereas the d.c. generator has a split rings commutator.
- (ii) a.c. generator produces current voltage in sinusoidal form but d.c. generator produces current flowing in one direction all through.
2. Four essential parts of a generator are armature, field magnet, slip rings and brushes.
3. The commutator converts a.c. wave form to d.c. wave form.
4. Attached to the bicycle for lighting purpose.

**ANSWERS TO TERMINAL EXERCISE**

1. (a)  $3 \times 10^{-2} \text{ W}_b \text{ s}^{-1}$                       (b)  $0.47 \text{ T s}^{-1}$
3. (a) becomes 4 times                      (b)  $0.8 \text{ H}$
8.  $333.3 \text{ V}$
10.  $0.2 \text{ Tesla}$
14.  $0.8 \text{ Volts}$



# ALTERNATING CURRENTS

## INTRODUCTION

In our household items, mostly we use the alternating current rather than direct current. The transmission of alternating current is easy compared to direct current. We can easily increase and decrease the alternating voltage or current with transformers. The current whose magnitude changes continuously and direction changes periodically, known as alternating current, whereas direct current has a constant magnitude and flows in only one direction (Fig. 23.1). In India, the frequency of alternating current is 50 Hz, i.e. it changes its direction 50 times in a second. The frequency of direct current is zero.

### OBJECTIVES

After studying this lesson, you should be able to

- explain alternating current and voltage;
- derive the relationship between voltage and current in ac circuits containing a (i) resistor (ii) inductor and (iii) capacitor;
- explain LC oscillations;
- describe the transformer construction and differentiate its types.

## 23.1) ALTERNATING CURRENT AND VOLTAGES

The alternating voltage varies like sine function with time as shown in Fig. 23.1. The current flowing in a circuit under these conditions is called alternating current.

In general, alternating voltage and currents are mathematically expressed as

$$V = V_m \cos \omega t \quad (23.1)$$

and 
$$I = I_m \cos \omega t \quad (23.2)$$

$V_m$  and  $I_m$  are known as the **peak values** of the alternating voltage and current respectively. In addition, we also define the root mean square (*rms*) values of  $V$  and  $I$  as

$$V_{rms} = \frac{V_m}{\sqrt{2}} = 0.707V_m \quad (23.3)$$

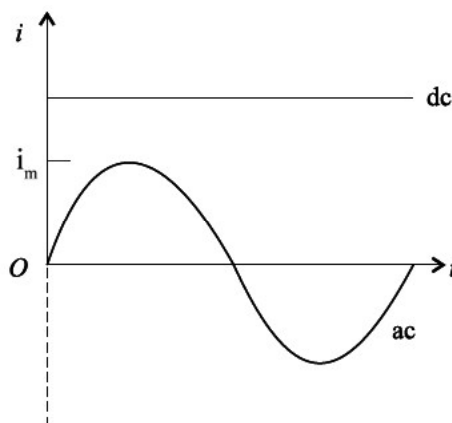


Fig. 23.1 : dc and ac current waveforms

$$I_{rms} = \frac{I_m}{\sqrt{2}} = 0.707I_m \quad (23.4)$$

The relation between  $V$  and  $I$  depends on the circuit elements present in the circuit. Let us now study a.c. circuits containing (i) a resistor (ii) a capacitor, and (iii) an inductor only.

### George Westinghouse (1846-1914)

If ac prevails over dc all over the world today, it is due to the vision and efforts of George Westinghouse. He was an American inventor and entrepreneur having about 400 patents to his credit. His first invention was made when he was only fifteen year old. He invented air brakes and automatic railway signals, which made railway traffic safe.



When Yugoslav inventor Nicole Tesla (1856-1943) presented the idea of rotating magnetic field, George Westinghouse immediately grasped the importance of his discovery. He invited Tesla to join him on very lucrative terms and started his electric company. The company shot into fame when he used the energy of Niagra falls to produce electricity and used it to light up a town situated at a distance of 20 km.

## 23.2) AC VOLTAGE APPLIED TO A RESISTOR

Refer to Fig. 23.2 which shows a resistor in an ac circuit. The instantaneous value of the current is given by the instantaneous value of the potential difference across the resistor divided by the resistance.

$$I = \frac{V}{R}$$

$$= \frac{V_m \cos \omega t}{R} \quad (23.5)$$

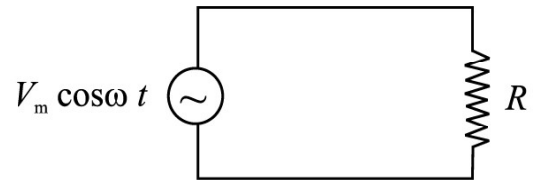


Fig. 23.2 : An ac circuit containing a resistor

The quantity  $V_m/R$  has units of volt per ohm, (i.e., ampere). It represents the maximum value of the current in the circuit. The current changes direction with time, and so we use positive and negative values of the current to represent the two possible current directions. Substituting  $I_m$ , the maximum current in the circuit, for  $V_m/R$  in Eqn. (23.5), we get

$$I = I_m \cos \omega t \quad (23.6)$$

Fig. 23.3 shows the time variation of the potential difference between the ends of a resistor and the current in the resistor. Note that the potential difference and current are in phase i.e., the peaks and valleys occur at the same time.

In India, we have  $V_m = 310\text{V}$  and  $\nu = 50\text{ Hz}$ . Therefore for  $R = 10\ \Omega$  we get

$$V = 310 \cos(2\pi 50t)$$

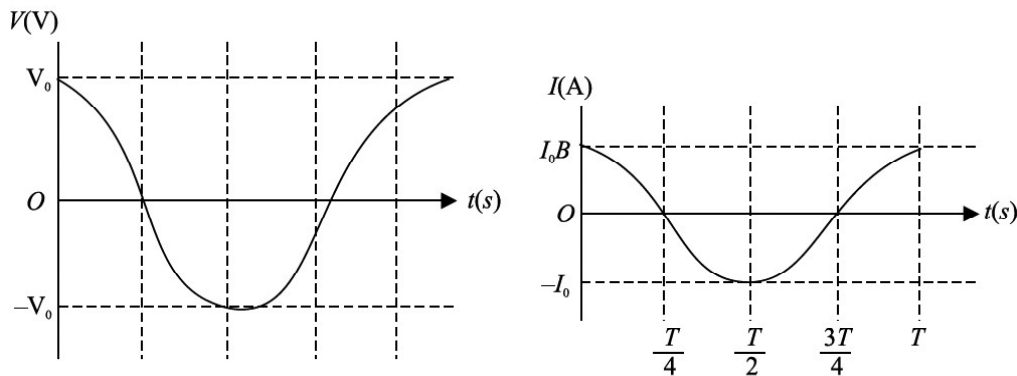


Fig. 23.3 : Time variation of current and voltage in a purely resistive circuit

and

$$I = \frac{310}{10} \cos(100\pi t)$$

$$= 31 \cos(100\pi t) \text{ A}$$

Since  $V$  and  $I$  are proportional to  $\cos(100\pi t)$ , the average current is zero over an integral number of cycles.

The **average power**,  $P = I^2 R$  developed in the resistor is not zero, because square of instantaneous value of current is always positive. As  $I^2$ , varies periodically between zero and  $I^2$ , we can determine the average power,  $P$ , for single cycle:

$$P_{\text{av}} = (I^2 R) = R (I^2)_{\text{av}} = R \frac{I_m^2 + 0}{2}$$

$$P_{\text{av}} = R \left( \frac{I_m^2}{2} \right) = R I_{\text{rms}}^2 \quad (23.7)$$

Note that the same power would be produced by a constant *dc* current of value  $\frac{I_m}{\sqrt{2}}$  in the resistor. It would also result if we were to connect the resistor to a potential difference having a constant value of  $\frac{V_m}{\sqrt{2}}$  volt. The quantities  $\frac{I_m}{\sqrt{2}}$  and  $\frac{V_m}{\sqrt{2}}$  are called the rms values of the current and potential difference. The term rms is short for root-mean-square, which means “the square root of the mean value of the square of the quantity of interest.” For an electric outlet in an Indian home where  $V_m = 310\text{V}$ , the rms value of the potential difference is

$$V_{\text{rms}} = \frac{V_m}{\sqrt{2}} \approx 220\text{V}$$

This is the value generally quoted for the potential difference. Note that when potential difference is 220 V, the peak value of a.c voltage is 310 V and that is why it is so fatal.

### Intext Questions 23.1

1. In a light bulb connected to an ac source the instantaneous current is zero two times in each cycle of the current. Why does the bulb not go off during these times of zero current?
2. An electric iron having a resistance 25 W is connected to a 220 V, 50 Hz household outlet. Determine the average current over the whole cycle, peak current, instantaneous current and the rms current in it.
3. Why is it necessary to calculate root mean square values of ac current and voltage.

## 23.3 REPRESENTATION OF AC CURRENT AND VOLTAGE BY ROTATING VECTORS - PHASORS

We have seen that the current through a resistor is in phase with voltage, But in case of inductor and capacitor, they are not in phase. In order to know the Phase relationships between voltage and current, we use the certain notation called Phasors. A phasor is a vector which rotates about its origin by an angular frequency  $\omega$ .

The vertical components of voltage and current at time  $t_1$  is  $V_m \sin \omega t_1$  and  $I_m \sin \omega t_1$  respectively.  $V_m$  and  $I_m$  are the peak values of voltage and current. The projection of voltage and current phasors on y-axis are  $V_m \sin \omega t$  and  $I_m \sin \omega t$  represent the value of voltage and current at that instant. If it rotated with frequency  $\omega$ , the curves are generated as shown in figure. The phasors  $V$  and  $I$  for the resistor are in the same direction as shown in Fig. 23.4 (a). Therefore the phase angle between the voltage and the current is zero. In case of inductor or capacitor, the voltage and currents are not in phase.

When an ac voltage is applied to a circuit consisting an inductor, the current in the circuit lags the voltage by  $\frac{\pi}{2}$  angle or one-quarter cycle. As shown in Fig. 23.4 (b) the current

phasor  $I$  is  $\frac{\pi}{2}$  behind the voltage phasor  $V$ .

When they rotated with frequency  $\omega$  counter clockwise, they generate voltage and current as shown in figure. From the figure, we see that current reaches its maximum value later than the voltage by one-fourth of a period.

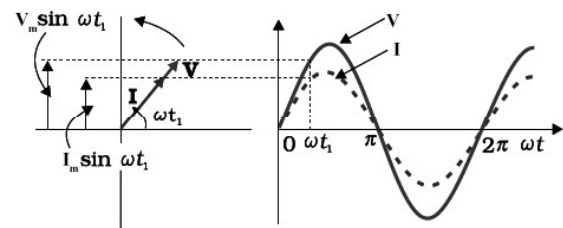


Fig. 23.4 : (a) Phasor diagram & wave diagram with resistor

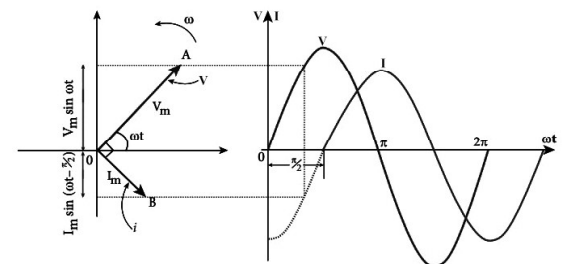


Fig. 23.4 : (b) Phasor diagram & wave diagram with inductor

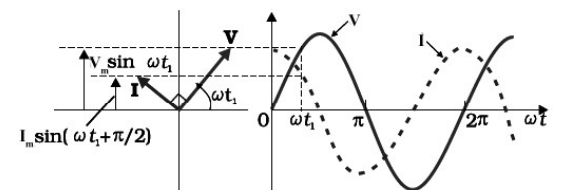


Fig. 23.4 : (c) Phasor diagram & wave diagram with capacitor

When an ac voltage is applied to a circuit consisting a capacitor, the current is  $\frac{\pi}{2}$  ahead of voltage. As shown in Fig. 23.4 (c), the current phasor  $\mathbf{I}$  is  $\frac{\pi}{2}$  ahead of the voltage phasor  $\mathbf{V}$  as they rotate counter clockwise. The variation of voltage and current with time is shown in figure and we observe that the current reaches its maximum value earlier than the voltage by one-fourth of a period.

### 23.4) AC VOLTAGE APPLIED TO A CAPACITOR

Fig. 23.5 shows a capacitor connected to an ac source. From the definition of capacitance, it follows that the instantaneous charge on the capacitor equals the instantaneous potential difference across it multiplied by the capacitance,  $q = CV$ . Thus, we can write

$$q = CV_m \cos \omega t \quad (23.8)$$

Since  $I = dq/dt$ , we can write

$$I = -\omega CV_m \sin \omega t \quad (23.9)$$

Time variation of  $V$  and  $I$  in a capacitive circuit is shown in Fig. 23.6.

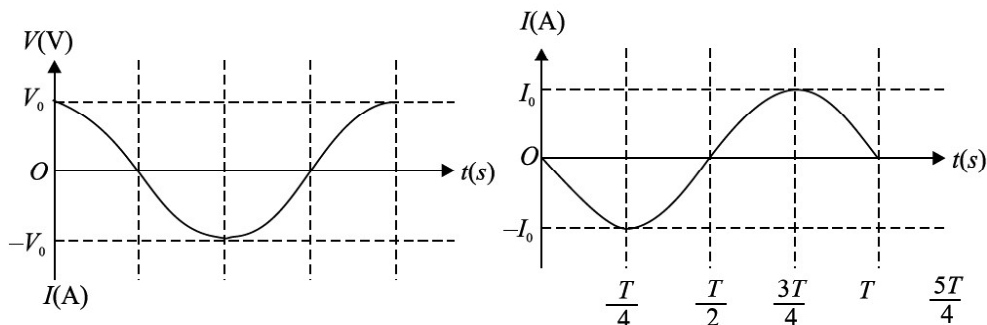


Fig. 23.5 : Capacitor in an ac circuit

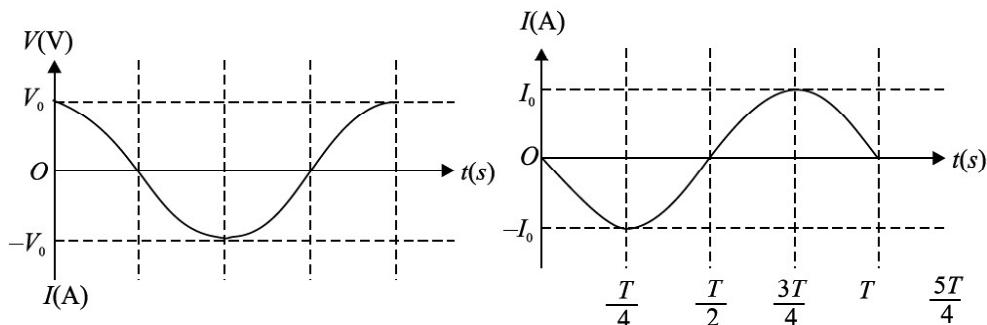


Fig. 23.6 : Variation in  $V$  and  $I$  with time in a capacitive circuit

Unlike a resistor, the current  $I$  and potential difference  $V$  for a capacitor are not in phase. The first peak of the current-time plot occurs one quarter of a cycle before the first peak in the potential difference-time plot. Hence we say that the capacitor current leads capacitor potential difference by one quarter of a period. One quarter of a period corresponds to a phase difference of  $\pi/2$  or  $90^\circ$ . Accordingly, we also say that the potential difference lags the current by  $90^\circ$ .

Rewriting Eqn. (23.9) as

$$I = -\frac{V_m}{1/(\omega C)} \sin \omega t \quad (23.10)$$

and comparing Eqns. (23.5) and (23.10), we note that  $(1/\omega C)$  must have units of resistance. The quantity  $1/\omega C$  is called the capacitive reactance, and is denoted by the symbol  $X_C$ :

$$X_C = \frac{1}{\omega C}$$

$$= \frac{1}{2\pi\nu C} \quad (23.11)$$

Capacitive reactance is a measure of the extent to which the capacitor limits the ac current in the circuit. It depends on capacitance and the frequency of the generator. The capacitive reactance decreases with increase in frequency and capacitance. Resistance and capacitive reactance are similar in the sense that both measure limitations to ac current. But unlike resistance, capacitive reactance depends on the frequency of the ac (Fig. 23.7). The concept of capacitive reactance allows us to introduce an equation analogous to the equation  $I = V/R$  :

$$I_{\text{rms}} = \frac{V_{\text{rms}}}{X_C} \quad (23.12)$$

The instantaneous power delivered to the capacitor is the product of the instantaneous capacitor current and the potential difference :

$$P = VI$$

$$= -\omega CV^2 \sin \omega t \cos \omega t$$

$$= -\frac{1}{2} \omega CV^2 \sin 2\omega t \quad (23.13)$$

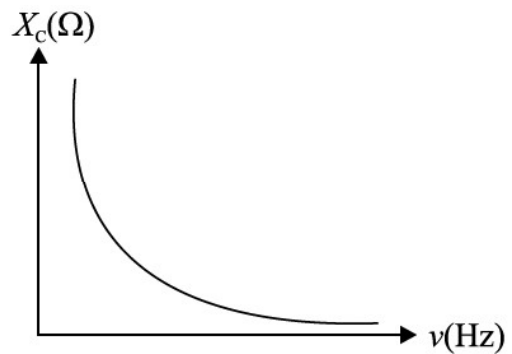


Fig. 23.7 : Frequency variation of capacitive reactance

The sign of  $P$  determines the direction of energy flow with time. When  $P$  is positive, energy is stored in the capacitor. When  $P$  is negative, energy is released by the capacitor. Graphical representations of  $V$ ,  $I$ , and  $P$  are shown in Fig. 23.8. Note that

whereas both the current and the potential difference vary with angular frequency  $\omega$ , the power varies with angular frequency  $2\omega$ . The average power is zero. The electric energy stored in the capacitor during a charging cycle is completely recovered when the capacitor is discharged. On an average, there is no energy stored or lost in the capacitor in a cycle.

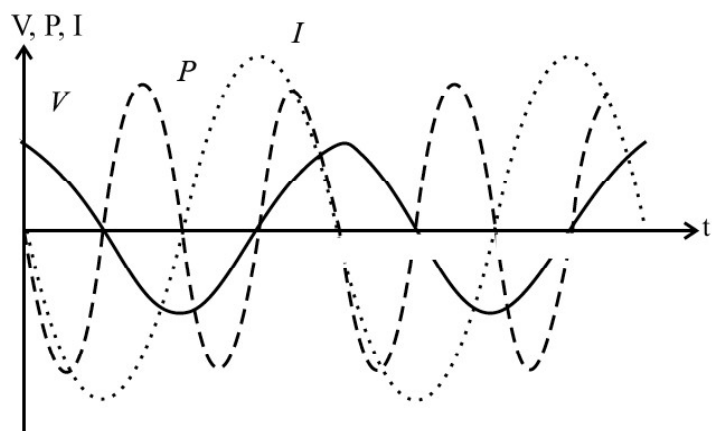


Fig. 23.8 : Time variation of  $V$ ,  $I$  and  $P$

**Example 23.1**

A  $100 \mu\text{F}$  capacitor is connected to a  $50 \text{ Hz}$  ac generator having a peak amplitude of  $220\text{V}$ . Calculate the current that will be recorded by an rms ac ammeter connected in series with the capacitor.

**Solution :**

The capacitive reactance of a capacitor is given by

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi(50 \text{ rads}^{-1})(100 \times 10^{-6} \text{ F})} = 31.8\Omega$$

Assuming that ammeter does not influence the value of current because of its low resistance, the instantaneous current in the capacitor is given by

$$\begin{aligned} I &= \frac{V}{X_C} \cos \omega t = \frac{220}{31.8} \cos \omega t \\ &= (-6.92 \cos \omega t) \text{ A} \end{aligned}$$

The rms value of current is

$$\begin{aligned} I_{\text{rms}} &= \frac{I_m}{\sqrt{2}} \\ &= \frac{6.92}{\sqrt{2}} \\ &= 4.91 \text{ A} \end{aligned}$$

Now answer the following questions.

**Intext Questions 23.2**

1. Explain why current in a capacitor connected to an ac generator increases with capacitance.
2. A capacitor is connected to an ac generator having a fixed peak value ( $V_m$ ) but variable frequency. Will you expect the current to increase as the frequency decreases?
3. Will average power delivered to a capacitor by an ac generator to be zero? Justify your answer.
4. Why do capacitive reactances become small in high frequency circuits, such as those in a TV set?



### 23.5 AC SOURCE CONNECTED TO AN INDUCTOR

We now consider an ideal (zero-resistance) inductor connected to an ac source. (Fig. 23.9). If  $V$  is the potential difference across the inductor, we can write

$$V(t) = L \frac{dI(t)}{dt} = V_m \cos \omega t \quad (23.14)$$

To integrate Eqn. (23.14) with time, we rewrite it as

$$\int dI = \frac{V_m}{L} \int \cos \omega t dt$$

Since integral of  $\cos x$  is  $\sin x$ , we get

$$I(t) = \frac{V_m}{\omega L} \sin \omega t + \text{constant} \quad (23.15)$$

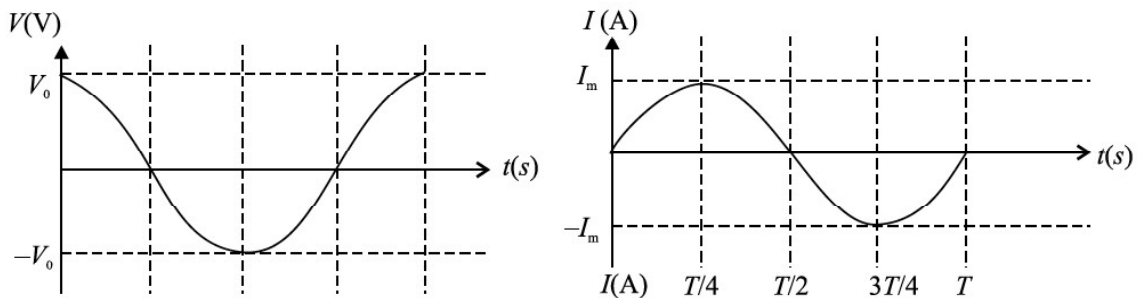
When  $t = 0$ ,  $I = 0$ . Hence constant of integration becomes zero. Thus

$$I(t) = \frac{V_m}{\omega L} \sin \omega t \quad (23.16)$$

To compare  $V(t)$  and  $I(t)$  let us take  $V_m = 220\text{V}$ ,  $\omega = 2\pi(50)\text{ rad s}^{-1}$ , and  $L = 1\text{ H}$ . Then

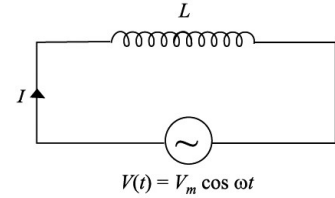
$$V(t) = 220 \cos(2\pi 50t) \text{ volt}$$

$$I(t) = \frac{220}{2\pi 50} \sin(2\pi 50t) = 0.701 \sin(2\pi 50t) \text{ ampere}$$



**Fig. 23.10 :** Time variation of the potential difference across an inductor and the current flowing through it. These are not in phase

Fig. 23.10 shows time variation of  $V$  and  $I$ . The inductor current and potential difference across it are not in phase. In fact the potential difference peaks one-quarter cycle before the current. We say that in case of an inductor current lags the potential difference by  $\pi/2$  rad (or  $90^\circ$ ). This is what we would expect from Lenz's law. Another way of seeing this is to rewrite Eqn. (23.16) as



**Fig. 23.9 :** An ac generator connected to an inductor

$$I = \frac{V_m}{\omega L} \cos\left(\omega t - \frac{\pi}{2}\right)$$

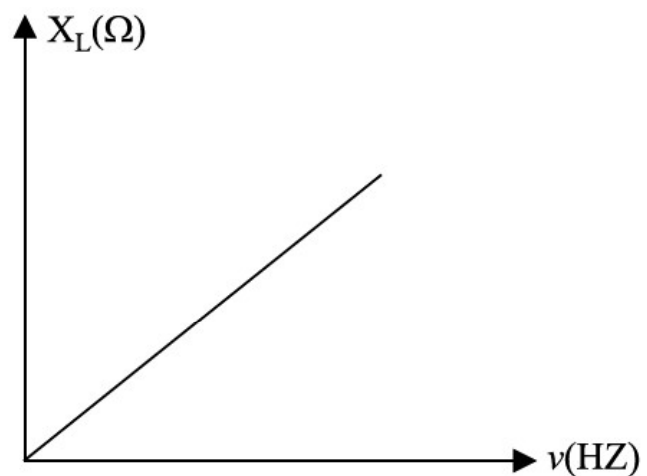
Because  $V = V_m \cos \omega t$ , the phase difference  $(-\pi/2)$  for  $I$  means that current lags voltage by  $\pi/2$ . This is in contrast to the current in a capacitor, which leads the potential difference. For an inductor, the current lags the potential difference.

The quantity  $\omega L$  in Eqn. (23.16) has units of resistance and is called **inductive reactance**.

It is denoted by symbol  $X_L$  :

$$X_L = \omega L = 2\pi\nu L \quad (23.17)$$

Like capacitive reactance, the inductive reactance,  $X_L$ , is expressed in ohm. **Inductive reactance** is a measure of the extent to which the inductor limits ac current in the circuit. It depends on the inductance and the frequency of the generator. Inductive reactance increases, if either frequency or inductance increases. (This is just the opposite of capacitive reactance.) In the limit frequency goes to zero, the inductive reactance goes to zero. But recall that as  $\omega \rightarrow 0$ , capacitive reactance tends to infinity (see Table-23.1). Because inductive effects vanish for a dc source, such as a battery, zero inductive reactance for zero frequency is consistent with the behaviour of an inductor connected to a dc source. The frequency variation of  $X_L$  is shown in Fig. 23.11.



**Fig. 23.11** : The reactance of an inductor ( $X_L = 2\pi\nu L$ ) as a function of frequency. The inductive reactance increases as the frequency increases.

**Table - 23.1** : Frequency response of passive circuit elements

Circuit element	Opposition to flow of current	Value at low-frequency	Value at high-frequency
Resistor	$R$	$R$	$R$
Capacitor	$X_C = \frac{1}{\omega C}$	$\infty$	$0$
Inductor	$X_L = \omega L$	$0$	$\infty$

The concept of inductive reactance allows us to introduce an inductor analog in the equation  $I = V/R$  involving resistance  $R$  :

$$I_{\text{rms}} = \frac{V_{\text{rms}}}{X_L} \quad (23.18)$$

The instantaneous power delivered to the inductor is given by

$$P = VI$$

$$= \frac{V_m}{\omega L} \sin \omega t \cos \omega t = \frac{V_m^2}{2\omega L} \sin 2\omega t \quad (23.19)$$

Graphical representations of  $V$ ,  $I$  and  $P$  for an inductor are shown in Fig. 23.12. Although both the current and the potential difference vary with angular frequency, the power varies with twice the angular frequency. The average power delivered over a whole cycle is zero. Energy is alternately stored and released as the magnetic field alternately grows and winds. decays

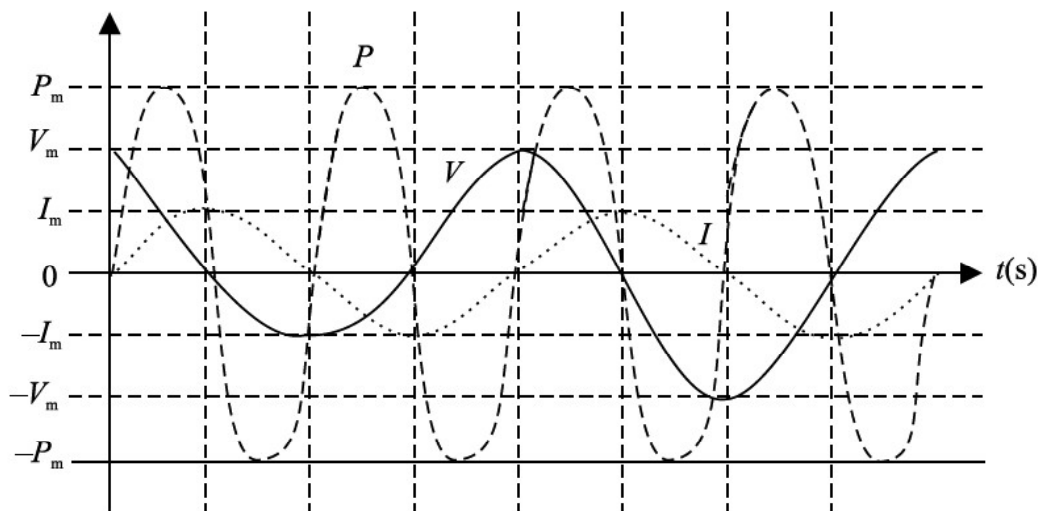


Fig. 23.12 : Time variation of potential difference, current and power in an inductive circuit

### Example 23.2

An air cored solenoid has length of 25 cm and diameter of 2.5 cm, and contains 1000 closely wound turns. The resistance of the coil is measured to be  $1.00 \Omega$ . Compare the inductive reactance at 100 Hz with the resistance of the coil.

#### Solution :

The inductance of a solenoid, whose length is large compared to its diameter, is given by

$$L = \frac{\mu_0 N^2 \pi a^2}{l}$$

where  $N$  denotes number of turns,  $a$  is radius, and  $l$  is length of the solenoid. On substituting the given values, we get

$$L = \frac{(4\pi \times 10^{-7}) \text{Hm}^{-1} (1000)^2 \pi (0.0125)^2 \text{m}^2}{0.25 \text{m}} = 2.47 \times 10^{-3} \text{H}$$

The inductive reactance at a frequency of 100 Hz is

$$X_L = \omega L = 2\pi \left( 100 \frac{\text{rad}}{\text{s}} \right) (2.47 \times 10^{-3}) \text{ H}$$

$$= 1.55 \Omega$$

Thus, inductive reactance of this solenoid at 100 Hz is comparable to the intrinsic (ohmic) resistance  $R$ . In a circuit diagram, it would be shown as

$$L = 2.47 \text{ H and } R = 1.00 \Omega$$


You may now like to test your understanding of these ideas.

### Intext Questions 23.3

- Describe the role of Lenz's law when an ideal inductor is connected to an ac generator.
- In section 23.2, self-inductance was characterised as electrical inertia. Using this as a guide, why would you expect current in an inductor connected to an ac generator to decrease as the self-inductance increases?

## 23.6) AC VOLTAGE APPLIED TO A SERIES LCR CIRCUIT

Refer to Fig. 23.13. It shows a circuit having an inductor  $L$ , a capacitor  $C$  and a resistor  $R$  in series with an ac source, providing instantaneous emf  $E = E_m \sin \omega t$ . The current through all the three circuit elements is the same in amplitude and phase but potential differences across each of them, as discussed earlier, are not in the same phase. Note that

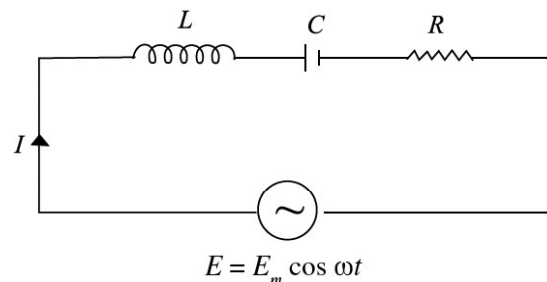


Fig. 23.13 : A series LCR circuit

- The potential difference across the resistor  $V_R = I_0 R$  and it will be in-phase with current.
- Amplitude of P.D. across the capacitor  $V_C = I_0 X_C$  and it lags behind the current by an angle  $\pi/2$  and
- amplitude of P.D. across the inductor  $V_L = I_0 X_L$  and it leads the current by an angle  $\pi/2$ .

Due to different phases, we cannot add voltages algebraically to obtain the resultant peak voltage across the circuit. To add up these voltages, we draw a phasor diagram showing proper phase relationship of the three voltages Fig. 23.14. The diagram clearly shows that voltages across the inductor and capacitor are in opposite phase and hence net voltage across the reactive components is  $(V_L - V_C)$ . The resultant peak voltage across the circuit is therefore given by

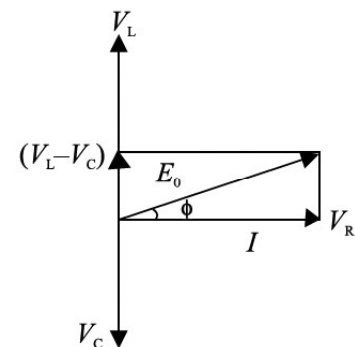


Fig. 23.14 : Phasor diagram of voltages across LCR.

$$E_0 = \sqrt{(V_L - V_C)^2 + V_R^2}$$

$$= \sqrt{I_0^2 \{(X_L - X_C)^2 + R^2\}}$$

or

$$\frac{E_0}{I_0} = \sqrt{(X_L - X_C)^2 + R^2}$$

The opposition to flow of current offered by a  $LCR$  circuit is called its *impedance*. The impedance of the circuit is given by

$$Z = \frac{E_{\text{rms}}}{I_{\text{rms}}} = \frac{E_0}{I_0} = \sqrt{(X_L - X_C)^2 + R^2} = \sqrt{\left(2\pi\nu L - \frac{1}{2\pi\nu C}\right)^2 + R^2} \quad (23.20)$$

Hence, the rms current across an  $LCR$  circuit is given by

$$I_{\text{rms}} = \frac{E_{\text{rms}}}{Z}$$

Also from Fig. 23.14 it is clear that in  $LCR$  circuit, the emf leads (or lags) the current by an angle  $\phi$ , given by

$$\tan \phi = \frac{V_L - V_C}{V_R} = \frac{X_L I_0 - X_C I_0}{R I_0} = \frac{X_L - X_C}{R} \quad (23.21)$$

This means that  $R$ ,  $X_L$ ,  $X_C$  and  $Z$  can also be represented on a phasor diagram similar to voltage (Fig. 23.15).

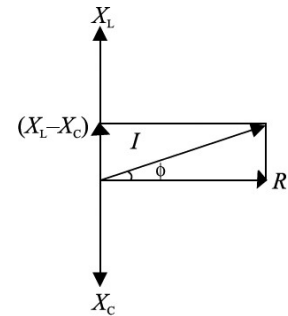


Fig. 23.15 : Phasor diagram for  $Z$

## Resonance

You now know that inductive reactance ( $X_L$ ) increases and capacitive reactance ( $X_C$ ) decreases with increase in frequency of the applied ac source. Moreover, these are out of phase. Therefore, there may be a certain frequency  $\nu_r$  for which  $X_L = X_C$  :

$$\text{i.e.,} \quad 2\pi\nu_r L = \frac{1}{2\pi\nu_r C}$$

$$\Rightarrow \quad \nu_r = \frac{1}{2\pi\sqrt{LC}} \quad (23.22)$$

This frequency is called *resonance frequency* and at this frequency, impedance has minimum value :  $Z_{\text{min}} = R$ . The circuit now becomes purely resistive. Voltage across the capacitor and the inductor, being equal in magnitude, annul each other. Since a resonant circuit is purely resistive, the net voltage is in phase with current ( $\phi = 0$ ) and maximum current flows through the circuit. The circuit is said to be in resonance with applied ac. The graphs given in Fig. 23.16 show the variation of peak value of current in an  $LCR$  circuit

with the variation of the frequency of the applied source. The resonance frequency of a given  $LCR$  circuit is independent of resistance. But as shown in Fig. 23.16, the peak value of current increases as resistance decreases.

The phenomenon of resonance in  $LCR$  circuits is utilised to tune our radio/TV receivers to the frequencies transmitted by different stations. The tuner has an inductor and a variable capacitor. We can change the natural frequency of the  $L-C$  circuit by changing the capacitance of the capacitor. When natural frequency of the tuner circuit

matches the frequency of the transmitter, the intercepting radio waves induce maximum current in our receiving antenna and we say that particular radio/TV station is tuned to it.

### Power in a LCR Circuit

You know that a capacitor connected to an ac source reversibly stores and releases electric energy. There is no net energy delivered by the source. Similarly, an inductor connected to an ac source reversibly stores and releases magnetic energy. There is no net energy delivered by the source. However, an ac generator delivers a net amount of energy when connected to a resistor. Hence, when a resistor, an inductor and a capacitor are connected in series with an ac source, it is still only the resistor that causes net energy transfer. We can confirm this by calculating the power delivered by the source, which could be a generator.

The instantaneous power is the product of the voltage and the current drawn from the source. Therefore, we can write

$$P = VI$$

On substituting for  $V$  and  $I$ , we get

$$\begin{aligned} P &= V_m \cos \omega t \left[ \frac{V_m}{Z} \cos(\omega t + \phi) \right] \\ &= \frac{V_m^2}{Z} \frac{2 \cos \omega t \cos(\omega t + \phi)}{2} \\ &= \frac{V_m^2}{2Z} \left[ \cos \phi + \cos \left( \omega t + \frac{\phi}{2} \right) \right] \end{aligned} \quad (23.23)$$

The phase angle  $\phi$  and angular frequency  $\omega$  play important role in the power delivered by the source. If the impedance  $Z$  is large at a particular angular frequency, the power will be small at all times. This result is consistent with the idea that impedance measures how the combination of elements impedes (or limits) ac current. Since the average value of the

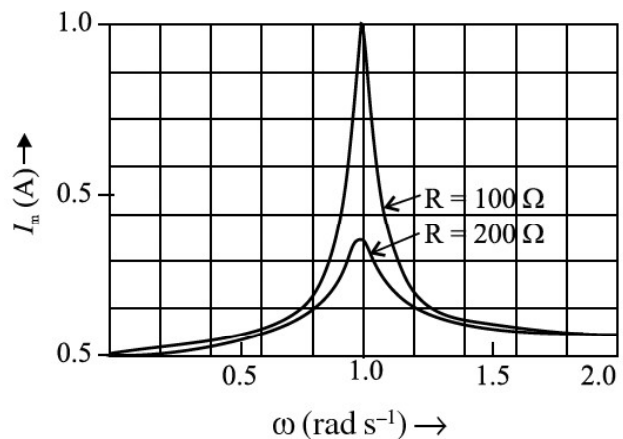


Fig. 23.16 : Variation of peak current in a LCR circuit with frequency for  
(i)  $R=100\Omega$ , and (ii)  $R=200\Omega$

second term over one cycle is zero, the average power delivered by the source to the circuit is given by

$$\text{Average Power} = \frac{V_m^2}{2Z} \cos \phi \quad (23.24)$$

$$= \frac{V_m}{\sqrt{2}} \cdot \frac{V_m}{\sqrt{2Z}} \cos \phi = V_{\text{rms}} I_{\text{rms}} \cos \phi \quad (23.25)$$

$\cos \phi$  is called *power factor* and is given by

$$\cos \phi = \frac{R}{Z} = \frac{R}{\sqrt{R^2 + (X_L - X_C)^2}} \quad (23.26)$$

The power factor delimits the maximum average power per cycle provided by the generator.

In a purely resistive circuit (or in a resonating circuit where  $X_L = X_C$ ),  $Z = R$ , so that  $\cos \phi = \frac{R}{R} = 1$ .

That is, when  $\phi = 0$ , the average power dissipated per cycle is maximum:  $P_m = V_{\text{rms}} I_{\text{rms}}$ .

On the other hand, in a purely reactive circuit, i.e., when  $R = 0$ ,  $\cos \phi = 0$  or  $\phi = 90^\circ$  and the average power dissipated per cycle  $P = 0$ . That is, the current in a pure inductor or pure capacitor is maintained without any loss of power. Such a current, therefore, is called *wattless current*.

### 23.7) LC OSCILLATIONS

The charged capacitor stores the electrical energy and inductor stores the magnetic energy.

When a charged capacitor is connected to an inductor, the charge on capacitor and the current in the circuit exhibit the electric oscillations called LC Oscillations.

LC Oscillations are generated due to the continuous energy flow from the capacitor to the inductor. In this, the capacitor undergoes charging, discharging, charging and so on. The capacitor starts charging the inductor while discharging itself when a charged capacitor is connected to a de-energized inductor.

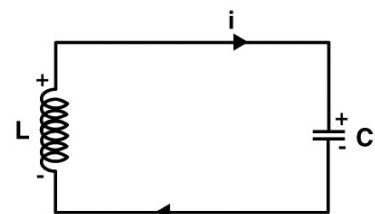


Fig. 23.17 :  
LC Oscillation Circuit

Let  $i$  be the current in the circuit at any instant of time  $t$  and  $q$  be the charge on the capacitor at that time. The potential difference across the capacitor is  $q/C$  and emf Induced in the emf is  $L di/dt$

Therefore the effective emf in the circuit is given by

$$\frac{q}{C} - L \frac{di}{dt} = 0 \quad (23.27)$$

$$\frac{d^2q}{dt^2} + \frac{1}{LC}q = 0 \quad \left( \because i = -\frac{dq}{dt} \right) \quad (23.28)$$

If we compare with the harmonic equation  $\frac{d^2q}{dt^2} + \omega_0^2q = 0$ , we can get the frequency

$$\omega_0 = \frac{1}{\sqrt{LC}} \quad (23.29)$$

and this will vary sinusoidally with time as  $q = q_m \cos(\omega_0 t + \phi)$  where  $q_m$  is the maximum value of  $q$  and  $\phi$  is a phase constant.

We know that 
$$i = -\frac{dq}{dt}$$

$$\therefore i = i_m \sin \omega_0 t, \text{ where } i_m = \omega_0 q_m$$

The LC oscillations are similar to mechanical oscillations like spring attached to a block.

**Table - 23.2 : The comparison of Mechanical oscillator with electrical oscillator**

Mechanical Oscillator	Electrical Oscillator
Mass (m)	Inductance (L)
Displacement (x)	Charge (q)
Velocity $\left( v = \frac{dx}{dt} \right)$	Current $\left( i = \frac{dq}{dt} \right)$
Mechanical Energy $E = \frac{1}{2}kx^2 + \frac{1}{2}mv^2$	Electromagnetic energy $U = \frac{1}{2} \frac{q^2}{C} + \frac{1}{2}Li^2$

## 23.8) TRANSFORMERS

Transformer is a device that changes (increases or decreases) the magnitude of alternating voltage or current based on the phenomenon of electromagnetic induction. A transformer has at least two windings of insulated copper wire linked by a common magnetic flux but the windings are electrically insulated from one another. The transformer windings connected to a supply source, which may be an ac main or the output of a generator, is called **primary winding**. The transformer winding connected to the load  $R_L$  is called the secondary winding. In the secondary winding, emf is induced when a.c. is applied to the primary. The primary and secondary windings, though electrically isolated from each other, are magnetically coupled with each other.



Basically, a transformer is a device which transfers electric energy (or power) from primary windings to secondary windings. The primary converts the changing electrical energy into magnetic energy. The secondary converts the magnetic energy back into electric energy.

An ideal transformer is one in which

- the resistance of the primary and secondary coils is zero;
- there is no flux leakage so that the same magnetic flux is linked with each turn of the primary and secondary coils; and
- there is no energy loss in the core.

Fig. 23.18 illustrates the configuration of a typical transformer. It consists of two coils, called primary and secondary, wound on a core (transformer). The coils, made of insulated copper wire, are wound around a ring of iron made of isolated laminated sheets instead of a solid core. The laminations minimize eddy currents in iron. Energy loss in a transformer can be reduced by using the laminations of “soft” iron for the core and thick high conductivity wires for the primary and secondary windings.

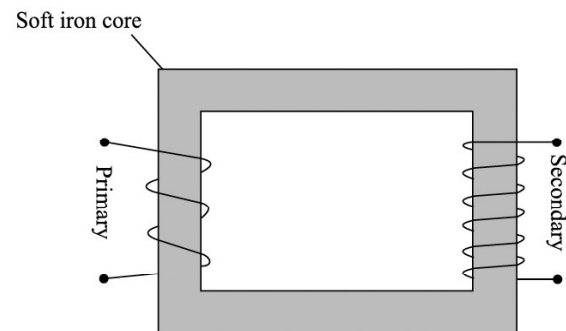


Fig. 23.18 : A schematic representation of a transformer

We now discuss the working of a transformer in the following two cases:

- (a) **Secondary an open circuit** : Suppose the current in the primary changes the flux through the core at the rate  $d\phi/dt$ . Then the induced (back) emf in the primary with  $N_p$  turns is given by

$$E_p = -N_p \frac{d\phi}{dt}$$

and the induced emf in the secondary windings of  $N_s$  turns is

$$\text{or} \quad E_s = -N_s \frac{d\phi}{dt} \quad ; \quad \frac{E_p}{E_s} = \frac{N_p}{N_s} \quad (23.30)$$

- (b) **Secondary not an open circuit** : Suppose a load resistance  $R_L$  is connected across the secondary, so that the secondary current is  $I_s$  and primary current is  $I_p$ . If there is no energy loss from the system, we can write

$$\text{Power input} = \text{Power output}$$

$$\text{or} \quad E_p I_p = E_s I_s$$

$$\text{so that} \quad \frac{I_p}{I_s} = \frac{E_s}{E_p} = \frac{N_p}{N_s} = k \quad (23.31)$$

Thus when the induced emf becomes  $k$  times the applied emf, the induced current is  $\frac{1}{k}$  times the original current. In other words, what is gained in voltage is lost in current.

### 23.8.1 Types of transformers

There are basically two types of transformers.

- (i) A **step-up transformer** increases the voltage (decreases the current) in secondary windings. In such transformers Fig. 23.19 (a) the number of turns in secondary is more than the number of turns in primary.
- (ii) A **step-down transformer** decreases the voltage (increases the current) in the secondary windings. In such transformers Fig. 23.19 (b), the number of turns in secondary is less than the number of turns in the primary.

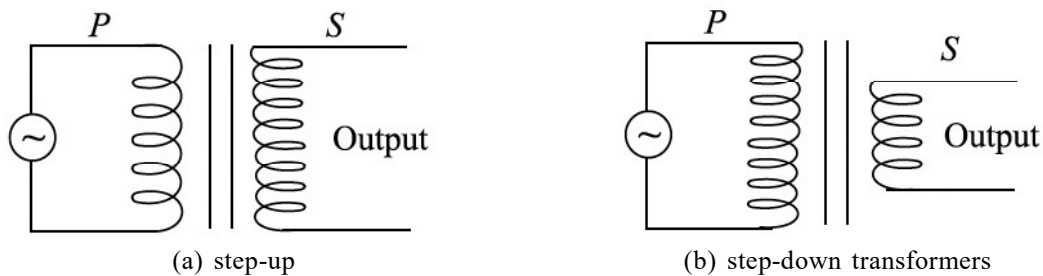


Fig. 23.19 : Iron cored

### 23.8.2 Efficiency of Transformers

While discussing the theory of the transformers we considered an ideal transformer in which there is no power loss. But in practice, some energy is always converted into heat in the core and the windings of the transformer. As a result, the electrical energy output across the secondary is less than the electrical energy input. The efficiency of a transformer is given by

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

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

The efficiency of a transformer is less than 100%. In a transformer the energy losses result from

- (a) Resistive heating in copper coils - *copper loss*,
- (b) Eddy current losses in form of heating of iron core - *Eddy current loss*.
- (c) Magnetization heating of the core during repeated reversal of magnetization - *hysteresis loss*.
- (d) Flux leakage from the *core*.

### Electrical Power Transmission

You have learnt how electricity is generated using ac or dc generators. You must have come across small units of generating sets in shops, offices and cinema halls. When power goes off, the mains is switched over to generator. In commercial use, generators which produce power of millions of watts at about 15 kV (kilo volt) is common. These generating plants may be hundreds of kilometers away from your town. Very large mechanical power (kinetic energy) is, therefore, necessary to rotate the rotor which produces magnetic field inside enormously large coils. The rotors are rotated by the turbines. These turbines are driven by different sources of energy.

To minimise loss of energy, power is transmitted at low current in the transmission lines. For this power companies step up voltage using transformers. At a power plant, potential difference is raised to about 330 kV. This is accompanied by small current. At the consumer end of the transmission lines, the potential difference is lowered using step down transformers.

You may now like to know how high potential difference used to transmit electrical power over long distances minimises current. We explain this with an example. Suppose electrical power  $P$  has to be delivered at a potential difference  $V$  by supply lines of total resistance  $R$ . The current  $I = P/V$  and the loss in the lines is  $I^2R = P^2R/V^2$ . It means that greater  $V$  ensures smaller loss. In fact, doubling  $V$  quarters the loss.

Electrical power is, thus, transmitted more economically at high potential difference. But this creates insulation problems and raises installation cost. In a 400kV super grid, currents of 2500 A are typical and the power loss is about 200 kW per kilometer of cable, i.e., 0.02% (percent) loss per kilometer. The ease and efficiency with which alternating potential differences are stepped-up and stepped-down in a transformer and the fact that alternators produces much higher potential difference than d.c. generators (25 kV compared with several thousand volts), are the main considerations influencing the use of high alternating rather than direct potential in most situations. However, due to poor efficiency and power thefts, as a nation, we lose about} Rs. 50,000 crore annually.

#### Example 23.3

What is the efficiency of a transformer in which the 1880 W of primary power provides for 1730 W of secondary power?

**Solution :**

Given  $P_{pri} = 1880\text{W}$  and  $P_{sec} = 1730\text{W}$ . Hence

$$\text{Efficiency} = \frac{P_{sec}}{P_{pri}} \times 100$$

$$\therefore = \frac{1730\text{W}}{1880\text{W}} \times 100 = 92\%$$

Thus, the transformer is 92% efficient.

### Example 23.4

A transformer has 100 turns in its primary winding and 500 turns in its secondary windings. If the primary voltage and current are respectively 120 V and 3 A, what are the secondary voltage and current?

**Solution :**

Given  $N_1 = 100$ ,  $N_2 = 500$ ,  $V_1 = 120\text{V}$  and  $I_1 = 3\text{A}$

$$V_2 = \frac{N_2}{N_1} \times V_1 = \frac{500 \text{ turns}}{100 \text{ turns}} \times 120\text{V} = 600\text{V}$$

$$I_2 = \frac{N_1}{N_2} \times I_1 = \frac{100 \text{ turns}}{500 \text{ turns}} \times 3\text{A} = 0.6\text{A}$$

### Intext Questions 23.4

1. Can a transformer work on dc? Justify your answer.
2. Why does step-up transformer have more turns in the secondary than in the primary?
3. Is the secondary to primary current ratio same as the secondary to primary voltage ratio in a transformer?
4. Toy trains often use a transformer to supply power for the trains and controls. Is this transformer step-up or step-down?

### WHAT YOU HAVE LEARNT

- In an LC circuit, the charge on the capacitor and the current in the circuit oscillate sinusoidally with the angular frequency  $\omega_0$  given by

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

- In an ac circuit, the voltage across the source is given by  $V = V_m \cos \omega t$  and current  $I = I_m \cos (\omega t + \phi)$
- In a purely resistive ac circuit, the voltage and current are in phase. The average power in such a circuit is  $P_{av} = \frac{I_m^2 R}{2}$

- In a purely capacitive ac circuit, the current leads the voltage by  $90^\circ$ . The average power in such a circuit is zero.
- In a purely inductive ac circuit, the current lags the voltage by  $90^\circ$ . The average power in such a circuit is zero.
- In a series LCR circuit, 
$$I_m = \frac{V_m}{Z} = \frac{V_m}{\left[ R^2 + (X_L - X_C)^2 \right]^{\frac{1}{2}}}$$

where  $Z$  is the impedance of circuit :  $Z = [R^2 + (X_L - X_C)^2]^{\frac{1}{2}}$

- For  $X_L - X_C = 0$ , an ac circuit is purely resistive and the maximum current  $I_m = \frac{V_m}{R}$ .

The circuit is said to be in resonance at  $\omega_0 = \frac{1}{\sqrt{LC}}$

- The average power  $P_{av} = V_{rms} \cdot I_{rms} = I_{rms}^2 R$ .
- A generator converts mechanical energy into electrical energy. It works on the principle of electromagnetic induction.
- A transformer is a static electrical device which converts an alternating high voltage to low alternating voltage and vice versa.
- The transformers are of two types: step-up to increase the voltage, and step-down : to decrease the voltage.
- The secondary to primary voltage ratio is in the same proportion as the secondary to primary turns ratio i.e.

$$\frac{V_s}{V_p} = \frac{N_s}{N_p}$$

- Main sources of power losses in a transformer are heating up of the windings and eddy current.
- For transmission of power from a power station to our homes, transformers and transmission lines are used.

### TERMINAL EXERCISE

- (a) Why is the rms value of an ac current always less than its peak value?
  - (b) The current in a  $2.5 \mu\text{F}$  capacitor connected to an ac source is given by  $I = -4.71 \sin 377t \mu\text{A}$   
Calculate the maximum voltage across the capacitor.

2. Calculate the capacitive reactance (for  $C = 2 \mu\text{F}$ ) and the inductive reactance (for  $L = 2 \text{ mH}$ ) at (i) 25 Hz and (ii) 50 Hz.
3. A series  $LCR$  circuit with  $R = 580 \Omega$ ,  $L = 31 \text{ mH}$ , and  $C = 47 \text{ nF}$  is driven by an ac source. The amplitude and angular frequency of the source are  $65 \text{ V}$  and  $33 \text{ krad/s}$ . Determine (a) the reactance of the capacitor, (b) the reactance of the inductor, (c) the impedance of the circuit, (d) the phase difference between the voltage across the source and the current, and (e) the current amplitude. Does current lead behind or lag the voltage across the source?
4. Explain why the reactance of a capacitor decreases with increasing frequency, whereas the reactance of an inductor increases with increasing frequency?
5. What is impedance of an  $LCR$  series circuit? Derive an expression for power dissipated in a.c.  $LCR$  circuit.
6. If the terminals of a battery are connected to the primary winding of transformer, why will a steady potential differences not appear across the secondary windings.
7. The power supply for a picture tube in a colour television (TV) set typically requires  $15,000 \text{ V A.C.}$  How can this potential difference be provided if only  $230 \text{ V}$  are available at a household electric outlet?
8. Would two coils acts as transformer without an iron core? If so, why not omit the core to save money?
9. An ac source has a 10-volt out-put. A particular circuit requires only a  $2 \text{ V A.C.}$  input. How would you accomplish this? Explain.
10. A person has a single transformer with 50 turns on one part of the core and 500 turns on the other. Is this a step-up or a step-down transformer? Explain.
11. Some transformers have various terminals or “taps” on the secondary so that connecting to different tap puts different functions of the total number of secondary windings into a circuit? What is the advantage of this?
12. A transformer in an electric welding machine draws  $3 \text{ A}$  from a  $240 \text{ V A.C.}$  power line and delivers  $400 \text{ A}$ . What is the potential difference across the secondary of the transformer?
13. A  $240\text{-V, } 400\text{W}$  electric mixer is connected to a  $120\text{-V}$  power line through a transformer. What is the ratio of turns in the transformer? and How much current is drawn from the power line?
14. The primary of a step-up transformer having 125 turns is connected to a house lighting circuit of  $220 \text{ V}_{\text{ac}}$ . If the secondary is to deliver  $15,000 \text{ volts}$ , how many turns must it have?
15. The secondary of a step-down transformer has 25 turns of wire and primary is connected to a  $220 \text{ V}_{\text{ac}}$  line. If the secondary is to deliver  $2.5 \text{ volt}$  at the out-put terminals, how many turns should the primary have?

16. The primary of a step-down transformer has 600 turns and is connected to a 120 V<sub>ac</sub> line. If the secondary is to supply 5 volts at its terminal and electron current of 3.5 A find the number of turns in the secondary and the electron current in the primary?
17. A step-up transformer with 352 turns in the primary is connected to a 220 V ac line. The secondary delivers 10,000 volts at its terminal and a current of 40 milliamperes.
- How many turns are in the secondary?
  - What is the current in the primary?
  - What power is drawn from the line?

### ANSWERS TO INTEXT QUESTIONS

#### 23.1

- It actually does but we cannot detect it, because the frequency of our domestic ac is 50 Hz. Our eye cannot detect changes that take place faster than 15 times a second.
- $$I_{rms} = \frac{E_{rms}}{R} = \frac{220 V}{25 \Omega} = 8.8 A$$
  - Peak value of current  $I_m = \sqrt{2} I_{rms} = 1.4 \times 8.8 = 12.32 A$   
Instantaneous current =  $I_0 \sin 2\pi vt = 12.32 \sin 100\pi t$
  - Average value of current over integral number of cycles will be zero.
- Since an ac current varies sinusoidally, its average value over a complete cycle is zero but rms value is finite.

#### 23.2

- Capacitive reactance  $X_C = \frac{1}{2\pi\nu C}$  as  $C$  increases  $X_C$  decreases and  $I$  increases.
- A charged capacitor takes some time in getting discharged. As frequency of source increases it starts charging the capacitor before it is completely discharged. Thus the maximum charge on capacitor and hence maximum current flowing through the capacitor increases though  $V_m$  is constant.
- Because the energy stored in the capacitor during a charging half cycle is completely recovered during discharging half cycle. As a result energy stored in the capacitor per cycle is zero.
- Capacitive reactance  $X_C = \frac{1}{2\pi\nu C}$  as  $\nu$  increases  $X_C$  decreases. This is so because on capacitor plates now more charge accumulates.

## 23.3

- In accordance with Lenz's law a back emf is induced across the inductor when ac is passed through it. The back emf  $\varepsilon = -L \frac{dl}{dt}$ .
- $I_{rms} = \frac{R_{rms}}{X_L}$ , as frequency increases,  $X_L (= 2\pi\nu L)$  increases, hence  $I_{rms}$  decreases.

## 23.4

- No, because the working of a transformer depends on the principle of electromagnetic induction, which requires time varying current.
- Because the ratio of the voltage in primary and secondary coils is proportional to the ratio of number of their turns.
- No, they are reciprocal to each other.
- Step-down transformer.

<b>ANSWERS TO TERMINAL EXERCISE</b>
-------------------------------------

- (b)  $5 \times 10^{-2} \text{V}$
- Capacitive reactance (i)  $\frac{1}{\pi} \times 10^4 \Omega$  (ii)  $\frac{1}{2\pi} \times 10^4 \Omega$   
Inductive reactance (i)  $0.1\pi\Omega$  (ii)  $0.2\pi\Omega$
- (a)  $6.7 \times 10^2 \Omega$  (b)  $99 \Omega$  (c)  $813.9 \Omega$   
(d) 4 rad (e) 0.16 A (f) Current lags
- 1.8 V
- 1 : 2,  $\frac{10}{3} \text{A}$
- 8522 turns
- 2200 turns
- 25 turns, 0.14 A
- (a) 16000 turns (b)  $\frac{20}{11} \text{A}$  (c) 400W





# ELECTROMAGNETIC WAVES

## INTRODUCTION

In the previous chapter 23, we learned that an electric current generates a magnetic field and that two current-carrying wires exert a magnetic force on each other. Additionally, in chapter 22, we discussed that a changing magnetic field creates an electric field. Now, the question arises: can an electric field changing with time also give rise to a magnetic field? James Clerk Maxwell (1831-1879) argued that indeed, not only an electric current but also a time-varying electric field generates a magnetic field.

While investigating the magnetic field at a point outside a capacitor connected to a time-varying current, Maxwell observed an inconsistency in Ampere's circuital law. To address this, he proposed the existence of an additional current, which he termed the "displacement current," to resolve the inconsistency.

Maxwell formulated a set of equations that involved electric and magnetic fields, as well as their sources—the charge and current densities. These fundamental equations became known as Maxwell's equations. Alongside the Lorentz force formula, they mathematically express all the fundamental laws of electromagnetism. Together, these concepts paved the way for a comprehensive understanding of electromagnetic phenomena.

The most significant revelation resulting from Maxwell's equations is the existence of electromagnetic waves, which are interlinked time-varying electric and magnetic fields propagating through space. According to these equations, the speed of these waves was found to be remarkably close to the speed of light (approximately  $3 \times 10^8$  m/s), as determined from optical measurements. This ground-breaking discovery led to the astonishing conclusion that light itself is an electromagnetic wave. Maxwell's pioneering work successfully unified the realms of electricity, magnetism, and light. Subsequently, in 1885, Hertz experimentally confirmed the existence of electromagnetic waves. The technological applications of these waves by visionaries like Marconi revolutionized communication and shaped the modern era we experience today.

In this chapter, we will first explore the significance of displacement current and its consequences. Next, we will provide a comprehensive overview of electromagnetic waves, discussing their broad spectrum, spanning from  $\gamma$  rays to long radio waves. The methods employed for transmitting and receiving electromagnetic waves for communication purposes will be discussed in Chapter-29.

## OBJECTIVES

After studying this lesson, you should be able to

- understand the displacement current;
- explain electromagnetic waves and electromagnetic spectrum;
- distinguish different types of electromagnetic waves and their uses.

## 24.1 DISPLACEMENT CURRENT

We know that a current in conductor produces a magnetic field. The quantitative relationship between the conduction current  $i$  and the magnetic field  $\vec{B}$  produced is given by Ampere's law which is stated as

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 i \quad (24.1)$$

Where  $\mu_0$  is permeability constant and  $i$  is the conduction current through any surface bounded by a closed linear path around which the line integral is taken.

**Displacement Current:** Maxwell postulated that it is not only the current in a conductor that produces a magnetic field, but a changing electric field in a vacuum or dielectric, also produces a magnetic field. Thus, a changing electric field is equivalent to a current which flows as long as the electric field is changing and produces the same magnetic effect as an ordinary conduction current. This is known as 'displacement current'. The displacement current concept leads to a modification of the Ampere's law. This inturn responsible for the success of the electromagnetic theory.

**Modified Ampere's Law:** Let us consider a circuit containing a source of emf, a resistor  $R$ , a capacitor  $C$  and a switch  $S$  Fig. 24.1 (a). When the switch is closed, a current flows and charge accumulates on the plates of the capacitor. The current decreases as the charge on the plates grows. When the capacitor is fully charged, there is no current. There is no actual flow of charge between the plates of the capacitor during charging.

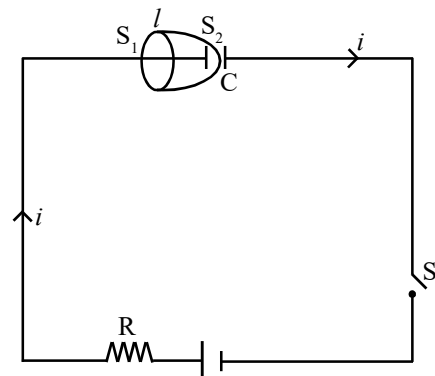


Fig. 24.1 (a)

Let us consider a plane surface  $S_1$  and a hemispherical surface  $S_2$ ; both bounded by the same closed path  $l$ .

During the charging process of Capacitor, current has been flowing through  $S_1$  but not through  $S_2$ . Therefore, if Ampere's law is applied to path  $l$  and the surface  $S_1$ , we find

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 i$$

When Ampere's law is applied to path  $l$  and surface  $S_2$ , we have  $\oint \vec{B} \cdot d\vec{l} = 0$

The above two equations cannot be correct. Since the contradiction arises from our use of Ampere's circuital law, this law must be missing something. The missing term must be such that one gets the same magnetic field for both the surfaces.

Maxwell removed this contradictory situation by adding one more term on the R.H.S. of Eqn. 24.1. It is for this purpose that he regarded the changing electric field in the gap between the plates of capacitor (during charging) as equivalent to a displacement current,  $i_d$ . The condition of continuity of the current requires the displacement current to be equal to  $\epsilon_0 \frac{d\phi_E}{dt}$ , where  $\epsilon_0$  is the permittivity of the free space and  $\phi_E$  is electric flux. Thus Ampere's law is modified to the form.

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 (i + i_d) = \mu_0 \left( i + \epsilon_0 \frac{d\phi_E}{dt} \right) \quad (24.2)$$

We can prove that displacement current  $i_d$  in the gap is equal to the conduction current  $i$ .

Let  $q$  be the charge on capacitor plates at an instant. Then the instantaneous electric field between the plates is

$$E = \frac{q}{\epsilon_0 A} \quad (24.3)$$

Where  $A$  is the surface area of each plate.

$$\therefore \frac{dE}{dt} = \frac{1}{\epsilon_0 A} \frac{dq}{dt} = \frac{i}{\epsilon_0 A} \quad (24.4)$$

Where  $i \left( = \frac{dq}{dt} \right)$  is the instantaneous current in the wires. Thus,  $i = \epsilon_0 A \frac{dE}{dt}$

By definition, the displacement current,  $i_d = \epsilon_0 \frac{d\phi_E}{dt}$

But,  $\phi_E = EA$

$$\therefore i_d = \epsilon_0 \frac{d}{dt} (EA) = \epsilon_0 A \frac{dE}{dt} \quad (24.5)$$

Comparing the equations for  $i$  and  $i_d$ , we find, that  $i_d = i$

This shows, that the displacement current in the gap is identical with the conduction current in the lead wires.

Thus the conception of displacement current makes the current continuous through the entire circuit including the capacitor Fig. 24.1 (b). The conduction current is not continuous across the capacitor gap because there is no actual flow of charge across the gap.

With this, Maxwell made the generalisation as that source of magnetic field is not just the conduction current but also the time rate of change of electric field.

Explicitly we can say, that outside the capacitor plates we have only conduction current ( $i$ ) and no displacement current ( $i_d$ ). On the other hand, inside the capacitor, there is no conduction current and there is only displacement current.

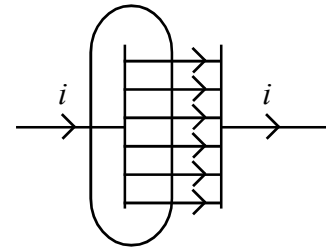


Fig. 24.1 (b)

Now, the generalised and correct Ampere's circuital law becomes as

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 i + \mu_0 \epsilon_0 \frac{d\phi_E}{dt} \quad (24.6)$$

and is known as Ampere - Maxwell law.

The consequence of displacement current is time dependent electric and magnetic fields give rise to each other.

### Maxwell's Equations

1.  $\oint \vec{E} \cdot d\vec{s} = \frac{q}{\epsilon_0}$  (Gauss law for electricity)
2.  $\oint \vec{B} \cdot d\vec{s} = 0$  (Gauss law for magnetism)
3.  $\oint \vec{E} \cdot d\vec{l} = -\frac{d\phi_B}{dt}$  (Faraday's law)
4.  $\oint \vec{B} \cdot d\vec{l} = \mu_0 i + \mu_0 \epsilon_0 \frac{d\phi_E}{dt}$  (Ampere - Maxwell Law)

The concept of displacement current is very important for insulators and for free space where the conduction current vanishes. For conductors, the displacement current is very small (less than 1% for frequencies less than  $10^{16}$  Hz) as compared to the conduction current and thus can be neglected.

## 24.2 ELECTROMAGNETIC WAVES

Stationary charges produces only electrostatic fields and charges in uniform motion produces only magnetic fields. But as result of Maxwell's theory that accelerated charges radiate electromagnetic waves.

Consider a charge in space which is oscillating with some frequency produces an oscillating electric field, which produces an oscillating magnetic field, which in turn is a source of oscillating electric field and so on. The oscillating electric and magnetic fields thus regenerate each other as the wave propagates through the space. The frequency of the electromagnetic wave equals to the frequency of oscillation of the charge. The energy associated with the propagating wave is the energy of the source i.e. the accelerated charge.

## Nature of Electromagnetic Wave

The light is an electromagnetic wave having different frequencies. Electric and magnetic fields vibrate perpendicular to each other and also perpendicular to direction of propagation.

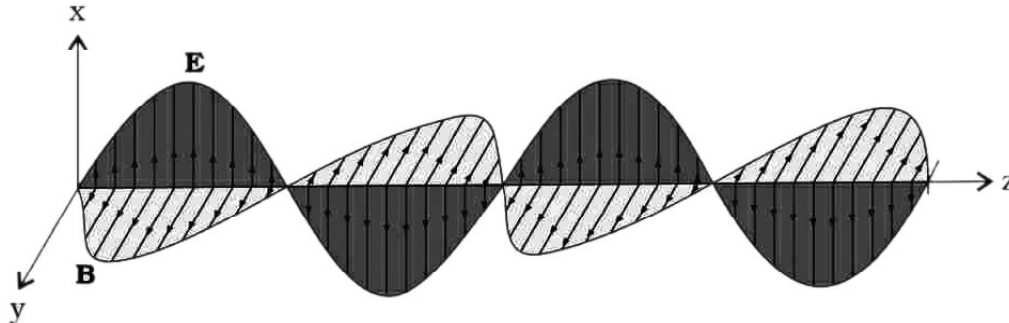


Fig. 24.2 : Propagation of Electromagnetic Waves

The Fig. 24.2 shows the typical electromagnetic wave propagating along the z-direction. The electric field  $E_x$  is along the x-axis and varies sinusoidally with z at a given time. The magnetic field  $B_y$  is along y-axis and varies sinusoidally with z.

The electric and magnetic fields  $E_x$  and  $B_y$  are perpendicular to each other and also perpendicular to z-axis which is a direction of propagation.

The electric field  $E_x$  is given by

$$E_x = E_0 \sin(kz - \omega t) \quad (24.7)$$

The magnetic field  $B_y$  is given by

$$B_y = B_0 \sin(kz - \omega t) \quad (24.8)$$

Where  $k$  is the magnitude of wave vector, which describe the direction of propagation.  $k$  is given by the equation

$$k = \frac{2\pi}{\lambda} \quad (24.9)$$

and  $\omega$  is the angular frequency. The speed of propagation of the wave is  $\frac{\omega}{k}$ .

Using the equation  $E_x$  and  $B_y$ , Maxwell's equations, we can find that  $\omega = ck$ , where  $c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$  which is the velocity of light.

From Maxwell's equations the relation between electric and magnetic field is given by

$$B_0 = \frac{E_0}{c} \quad (24.10)$$

### 24.2.1 Salient features of electromagnetic waves

- Electromagnetic waves self - sustaining oscillations of electric and magnetic fields in free space or vacuum.
- Speed of electromagnetic waves in a medium of permittivity  $\epsilon$  and magnetic permeability  $\mu$  is given by  $v = \frac{1}{\sqrt{\mu\epsilon}}$ .
- Electromagnetic waves can undergo diffraction, refraction and polarisation.
- Electromagnetic waves carry energy and momentum. In a region of free space with a electric field  $E$ , the energy density is  $\frac{\epsilon_0 E^2}{2}$  and energy density associated with magnetic field of  $B$  is  $\frac{B^2}{2\mu_0}$ .
- Electromagnetic waves exerts pressure as they carry energy and momentum.
- Electromagnetic waves carry energy from one place to another.

## 24.3 ELECTROMAGNETIC SPECTRUM

The position of an electromagnetic spectrum can be characterised by either its frequency or wavelength of oscillations. The classification of electromagnetic waves according to frequency is the electromagnetic spectrum. Depending on the frequency electromagnetic waves they are called with different names like radio waves, micro waves, infrared radiation, visible light, Ultraviolet radiation, x-rays and gamma rays. There is no sharp division between one kind of wave and the next wave. The classification is based roughly on how the waves are produced.

We briefly describe these different types of electromagnetic waves.

### 24.3.1 Radio Waves

The frequency range of Radio waves starts from 500 kHz to 1000 MHz. These are produced from accelerated motion of charges in conducting wires. Radio waves are used in Radio and Television communication systems. These waves are also used in shortwave radio, navigation, air - traffic control and cellular phones. In Radio communication the Amplitude Modulated (AM) band is from 530 kHz to 1710 kHz and the Frequency Modulated (FM) band is from 88 MHz to 108 MHz. In Television communication, the frequency range is from 54 MHz to 890 MHz. Higher frequencies upto 54 MHz are used in shortwave bands. In cellular phones, Ultra High Frequency (UHF) band is used for communication.

### 24.3.2 Microwaves

Microwaves are short wavelength radio waves. The frequency range of these waves is in the gigahertz (1 GHz to 1000 GHz) range. Microwaves are produced inside the oven by

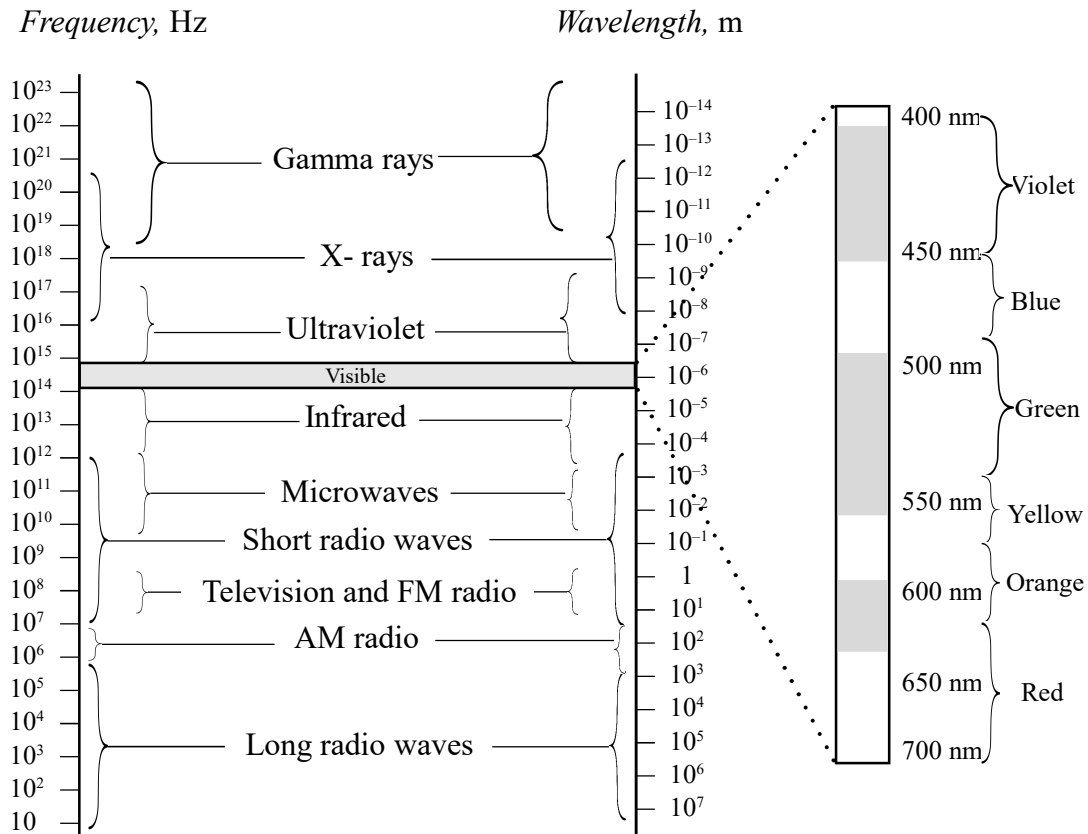


Fig. 24.3 : Electromagnetic Spectrum

an electron tube called a magnetron. These waves are also produced from special vacuum tubes called Klystrons and Gunn diodes.

The uses of the microwaves are similar to that of radio waves. These are used in communications, radio astronomy, radar systems used in aircraft navigation and microwave ovens for cooking of food. In microwave ovens, the frequency of microwaves are selected to match the resonant frequency of water molecules of food so that energy from the waves is transferred efficiently to the kinetic energy of the molecules. So the temperature of any food containing water raises, therefore it used to cook food or keep it warm.

### 24.3.3 Infrared Waves

The frequency of infrared waves ranges from 300 GHz to 400 THz and wavelength ( $\lambda$ ) between 780 nm to 1 mm. These are referred as heat waves because they have heat inducing property. Infrared waves are produced by hot bodies and molecules.

Infrared rays are classified into two categories i.e., Near - infrared and far - infrared.

Infrared rays are used in applications where heat production is required, example infrared heaters. Near infrared waves are used in electronic applications like TV remotes, video recorders and photography. These are also used in physical therapy as infrared lamps. Infrared radiation is very important to maintain the earth temperature warm through greenhouse effect. Infrared detectors are used in Earth satellites, both for military purposes and to observe growth of crops.

### 24.3.4 Visible rays

The visible rays are most familiar form of electromagnetic waves, which human eye can detect. The frequency range of visible rays is from  $4 \times 10^{14} \text{ Hz}$  to  $7 \times 10^{14} \text{ Hz}$  or wavelength range is 700 - 400 nm. The sun is the dominant source for visible light waves. Humans get the sensation of the objects and information of the world due to visible rays. Animals like snakes can detect infrared rays and insects can detect ultraviolet rays.

Visible rays are used in fibre optic communications, photography and in electronic devices.

### 24.3.5 Ultraviolet rays

The frequency of ultraviolet rays ranges from  $10^{14} \text{ Hz}$  to  $10^{16} \text{ Hz}$  and wavelength 400 nm to 0.6 nm. Ultraviolet rays are produced by special lamps and very hot bodies. The sun is main source of ultraviolet rays and most of the ultraviolet rays are absorbed by ozone layer. These rays are harmful to human body which causes skin diseases. Ultraviolet radiation can be protected by suitable glass windows. Ozone layer in the atmosphere plays a protective role and hence its depletion by chlorofluoro carbons is a serious international issue.

Ultraviolet rays are widely used in industrial processes and in medical applications such as killing bacteria, creating fluorescent effects, curing inks and resins, phototherapy and suntanning.

### 24.3.6 X-rays

The frequency range of X-rays is from  $10^{16}$  to  $10^{20} \text{ Hz}$  and wavelength is about  $10^{-8} \text{ m}$  (10 nm) to  $10^{-13} \text{ m}$  ( $10^{-4} \text{ nm}$ ). X-rays are produced when electrons are accelerated under a high potential difference and bombarded to a metal target. X-rays are used as diagnostic tool to detect fractures and breaks at the bones and joints. And also used to detect problems affecting soft tissues. X-rays are used to treat certain forms of cancer. Care must be taken, while using the x-rays because they damage living tissues and internal organs.

### 24.3.7 Gamma rays

The Gamma rays are high energy electromagnetic waves having high frequency more than  $10^{26} \text{ Hz}$  and wavelength is from  $10^{-10} \text{ m}$  to less than  $10^{-14} \text{ m}$ . Gamma rays are produced in nuclear reactions and also emitted by radioactive nuclei. These rays are also produced by the hottest and most energetic objects in the universe, such as neutron stars and pulsars, supernova explosions and regions around black holes.

Gamma rays are used in medicine for radiotherapy in industry for sterilization and disinfection and in the nuclear industry. Shielding against gamma rays is essential because they can cause diseases to skin or blood, eye disorders and cancers.



Table - 24.1: Frequency, Wavelength, Production and uses of different Electromagnetic rays

Sl. No.	Type of waves	Frequency range	Wavelength range	Production of waves	Uses
1.	Radio	500 KHz to 1000 MHz	> 0.1 m	accelerated motion of charges in conducting wires	Radio and television communication Air traffic control Cellular phones.
2.	Microwaves	1 GHz to 1000 GHz	0.1 m to 1 nm	by using magnetron, special vacuum tubes, klystrons and Gunn diodes.	Communications Radio astronomy Radar system Microwave ovens
3.	Infrared	300 GHz to 400 THz	780 nm to 1 nm	By hot bodies and molecules	Infrared heaters TV remotes, Photography Physical therapy
4.	Visible	$4 \times 10^{14}$ Hz to $7 \times 10^{14}$ Hz	700 - 400 nm	The sun is the dominant source for visible light	Fibre optic communications electronic devices.
5.	Ultraviolet	$10^{14}$ Hz to $10^{16}$ Hz	400 nm to 0.6 nm	Produced by special lamps and very hot bodies.	In industry and medical applications such as killing bacteria.
6.	X-rays	$10^{16}$ Hz to $10^{20}$ Hz	10 nm to $10^{-4}$ nm	When electrons are accelerated under a high potential difference and bombarded to a metal target.	In x-ray photographic film to detect fractures and breaks at the bones and joints. In the treatment of certain forms of cancer.
7.	Gamma	> $10^{20}$ Hz	$10^{-10}$ to $10^{-14}$ m	In nuclear reactors, neutron stars, pulsars, supernova explosions, black holes.	Medical applications In Industry In radio therapy

**Example 24.1**

Electromagnetic wave travels in free space along the  $x$ -direction. At a particular point in space and time,  $E = 9.6 \hat{j}$  V/m. Find the value  $B$  at that point.

**Solution :**

$$B = \frac{E}{c}$$

$$B = \frac{9.6 \text{ V/m}}{3 \times 10^8 \text{ m/s}} = 3.2 \times 10^{-8} \text{ T}$$

**Example 24.2**

Electromagnetic wave with an energy flux of  $12 \text{ w/cm}^2$  falls on a non reflecting surface at normal incidence. If the surface has an area of  $30 \text{ cm}^2$ , find the average force exerted on the surface during a 20 minute span.

**Solution :**

The total energy fall in on surface

$$U = 12 \times 30 \times 20 \times 60 \text{ J}$$

$$U = 4.32 \times 10^5 \text{ J}$$

Therefore, the total momentum

$$P = \frac{U}{c} = \frac{4.32 \times 10^5 \text{ J}}{3 \times 10^8 \text{ m/s}}$$

$$P = 1.44 \times 10^{-3} \text{ kg m/s}$$

The average force exerted on the surface

$$F = \frac{p}{t} = \frac{1.44 \times 10^{-3}}{1.2 \times 10^3}$$

$$F = 1.2 \times 10^{-6} \text{ N}$$

**INTEXT QUESTIONS**

1. Explain what type of radiation coming from Sun.
2. Explain displacement current.
3. Write the different types of Electromagnetic radiation.
4. What is the frequency and wavelength range of Ultraviolet rays?
5. How food was cooked with micro waves in microwave ovens?
6. What are the applications of Infrared rays?
7. The charging current for a parallel plate capacitor is 0.5 A. What is the displacement current across its plates?

**WHAT YOU HAVE LEARNT**

- The light is an electromagnetic wave having different frequencies.
- In electromagnetic wave, electric and magnetic fields vibrate perpendicular to each other and both are perpendicular to direction of propagation.

- Ampere's law in integral form  $\oint B \cdot dl = \mu_0 i$ .
- The displacement current  $i_d = \epsilon_0 \frac{d\phi_E}{dt}$ .
- The modified Ampere's law or Ampere's Maxwell's law  $\oint B \cdot dl = \mu_0 i + \mu_0 \epsilon_0 \frac{d\phi_E}{dt}$ .
- The electric field  $E_x$  and magnetic field  $B_y$  is given by  $E_x = E_0 \sin(kz - \omega t)$  and  $B_y = B_0 \sin(kz - \omega t)$  respectively.
- The magnitude of wave vector  $k = \frac{2\pi}{\lambda}$ .
- If  $\omega$  is the angular frequency then the speed of propagation of the wave is  $\frac{\omega}{k}$  and velocity of light is  $c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$ .
- The relation between electric and magnetic field is given by  $B_0 = \frac{E_0}{c}$ .
- Depending on frequency, electromagnetic waves are classified as radio waves, microwaves, infrared radiation, visible light, ultraviolet radiation, x-rays and gamma rays.

### TERMINAL EXERCISE

1. The magnetic field in a plane electromagnetic wave is given by  $B_y = 3 \times 10^{-7} \sin(0.4 \times 10^3 x + 1.6 \times 10^{11} t)$ . Then (i) Find the wavelength and frequency of the wave. (ii) Determine the expression for electric field.
2. The electric field in a plane electromagnetic wave is 900 N/C. The wave is propagating in the  $x$  - direction and the electric field in the  $y$  - direction. Determine the maximum magnetic field and its direction.
3. The electric field in a electromagnetic wave is given by  $E = 60 \sin \omega \left( t - \frac{x}{c} \right)$  N/C. Find the energy density. ( $\epsilon_0 = 8.85 \times 10^{-12} \text{ C/Nm}^2$ ).
4. The amplitude of the magnetic field part of a harmonic electromagnetic wave in vacuum is  $B_0 = 620 \text{ nT}$ . What is the amplitude of the electric field part of the wave?

5. The electric field amplitude of an electromagnetic wave is  $E_0 = 90 \text{ N/C}$  and that its frequency is  $40.0 \text{ MHz}$ . Find the  $B_0$  and  $\omega$ .
6. Find the amplitude of the electric field in a parallel beam of light of intensity  $3.0 \text{ W/m}^2$ . ( $\epsilon_0 = 8.85 \times 10^{-12} \text{ C/Nm}^2$ ).

### ANSWERS TO INTEXT QUESTIONS

1. The light coming from sun is a combination of electrical and magnetic components is called electromagnetic waves.
2. The changing electric field produces a magnetic field. So a changing electric field is equivalent to a current as long as electric field is changing and produces the same magnetic effect as an ordinary conduction current. This is called as displacement current.

The displacement current  $i_d = \epsilon_0 \frac{d\phi_E}{dt}$

3. The different types of electromagnetic radiations are Radio waves, microwaves, infrared, visible, ultraviolet, X-rays and Gamma rays.
4. The frequency of ultraviolet rays ranges from  $10^{14} \text{ Hz}$  to  $10^{16} \text{ Hz}$  and wavelength  $400$  to  $0.6 \text{ nm}$ .
5. The frequency of microwaves is selected to match the resonant frequency of water molecules of food so that energy from the wave is transferred efficiently to the kinetic energy of the molecules. So the temperature of any food containing water raises, therefore the food is cooked.
6. Applications of infrared rays are
  1. Infrared heaters
  2. TV remotes
  3. Photography
  4. Physical therapy
7. Charging current means conduction current and we know that conduction current is always equal to displacement current.

So, displacement current across its plates is  $0.5 \text{ A}$ .

<b>ANSWERS TO TERMINAL EXERCISE</b>
-------------------------------------

$$1. \quad \lambda = \frac{2\pi}{0.4 \times 10^3} = \frac{2 \times 3.4}{0.4 \times 10^3} = 1.57 \text{ cm}$$

$$E_0 = B_0 c = 3 \times 10^{-7} \times 3 \times 10^8 = 9 \times 10 = 90 \text{ V/m}$$

$$E_z = 90 \sin(0.4 \times 10^3 x + 1.6 \times 10^{11} t) \text{ V/m}$$

$$2. \quad E_0 = 900 \text{ N/C}$$

$$B_0 = \frac{E_0}{c} = \frac{900}{3 \times 10^8} = 3 \times 10^{-6} \text{ T}$$

Electric field is in y-direction and wave propagation is in x-direction, hence, direction of magnetic field is in z-direction.

$$3. \quad \text{Energy density} = \frac{1}{2} \epsilon_0 E_0^2$$

$$= \frac{1}{2} \times 8.85 \times 10^{-12} \times 60 \times 60$$

$$= 1.6 \times 10^{-8} \text{ J/m}^3$$

$$4. \quad E_0 = B_0 c = 620 \times 10^{-9} \times 3 \times 10^8 = 186 \text{ N/C}$$

$$5. \quad B_0 = \frac{E_0}{c} = \frac{90}{3 \times 10^8} = 3.0 \times 10^{-7} \text{ T}$$

$$\omega = 2\pi f = 2 \times \frac{22}{7} \times 40 \times 10^6 = 2.5 \times 10^8 \text{ Hz}$$

$$6. \quad E_0 = \sqrt{\frac{2I}{\epsilon_0 c}} = \sqrt{\frac{2 \times 3.0}{8.85 \times 10^{-12} \times 3 \times 10^8}} = \sqrt{0.225 \times 10^4}$$

$$= 47.3 \text{ N/C}$$



# DUAL NATURE OF RADIATION AND MATTER

## INTRODUCTION

You must have seen films in cinema halls. The picture on the screen is produced by passing light through films which have the scenes shot on them. But have you ever thought as to how the sound is reproduced in the cinema? The sound is also recorded on the side of the film as a sound track. The light beam passing through this sound track falls on a photocell, which converts it into electrical pulses. These electrical pulses are converted into sound. In this lesson you will study the effect which governs the working of a photocell. It is called the *photoelectric effect*. It is also used in burglar alarm to detect intruders. Einstein's explanation of photoelectric effect led de Broglie (read as de Brooy) to the wave-particle duality, i.e. matter exhibits wave as well as particle properties.

You now know that a particle is characterized by properties such as definite position, size, mass, velocity, momentum, etc. Its motion is described by Newton's laws of motion. On the other hand, a wave is characterized by properties such as periodicity in space-time, wavelength, amplitude, frequency, wave velocity, etc. It transports energy, but no matter. That is, it extends in space unlike a particle, which is localised. The term wave-particle duality refers to the behaviour where both wave-like and particle-like properties are exhibited under different conditions by the same entity. His arguments were simple: Nature likes simplicity and loves symmetry. So if wave-particle duality can be exhibited by light, it should be exhibited by matter as well. You will learn about his explanation of matter waves in sec. 25.3.

## OBJECTIVES

After studying this lesson, you should be able to

- *explain photoelectric effect;*
- *describe the experimental arrangement to study photoelectric effect;*
- *state the laws of photoelectric emission;*
- *interpret the graphs between frequency of radiation and retarding potential;*
- *write deBroglie wavelength of matter waves associated with a particle of momentum  $p$ ;*
- *describe the experimental arrangement for the verification of matter waves.*

## 25.1) PHOTOELECTRIC EFFECT

In 1887, while working on propagation of electromagnetic waves, Hertz discovered that air in a spark gap became a better conductor when it was illuminated by ultraviolet rays. Further experiments by him showed that zinc became positively charged on irradiation by ultraviolet rays. In 1900, Leonard showed that electrons were emitted from a metal surface when light of sufficiently high frequency falls on it. This phenomenon is known as *photoelectric effect* and the electrons so emitted are called photoelectrons.

*The emission of electrons from metals irradiated by light of a frequency greater than a certain characteristic frequency is called photoelectric effect.*

### 25.1.1 Experimental Arrangement to Study Photoelectric Effect

Refer to Fig. 25.1. It shows a schematic diagram of the apparatus that can be used to study this phenomenon.

A metallic plate (photo sensitive plate) *C* called photo cathode is sealed inside an evacuated tube along with another metal plate *A*, which is used to collect photoelectrons emitted by *C*. These electrodes are connected to a battery and micro-ammeter circuit, as shown in Fig. 25.1. The battery is so connected that the voltage on plate *A* is positive with respect to *C*. If the battery terminals are reversed, the voltage of the plate *A* will become negative relative to *C*.

The emission of electrons from metals can also take place when they are heated. This is known as **thermionic emission**. *Note that electrons gain energy from thermal energy in thermionic emission.*

To study the effect of intensity of incident light on the number of photoelectrons emitted by *C*, the collector plate *A* is kept at positive potential relative to *C*.

Keeping the frequency of incident light and the value of accelerating potential fixed, the photoelectrons emitted per unit area from the emitting surface vary linearly with the intensity of light, as shown in Fig. 25.2 (a).

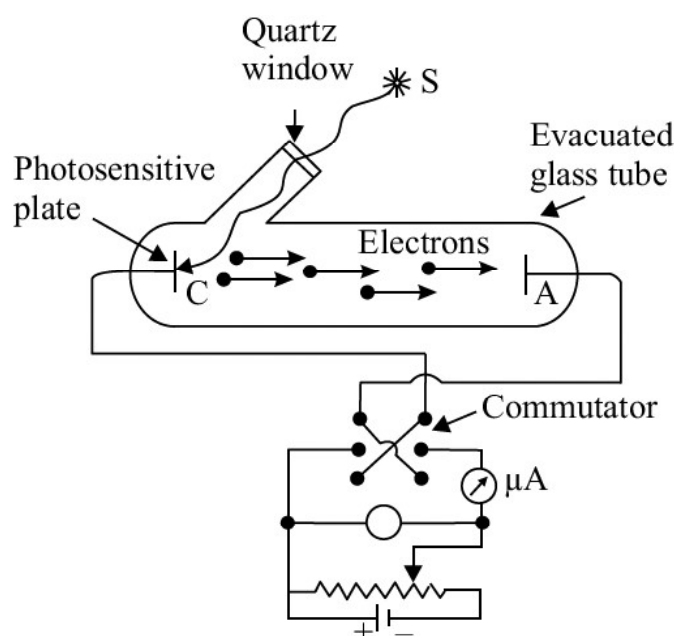


Fig. 25.1 : Experimental arrangement for observing the photoelectric effect

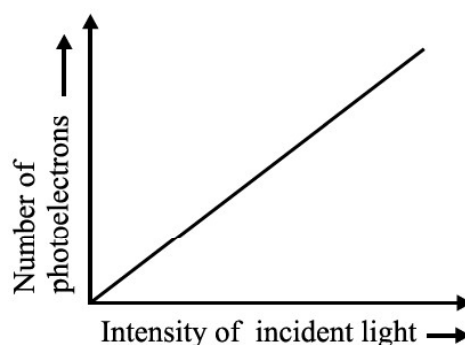
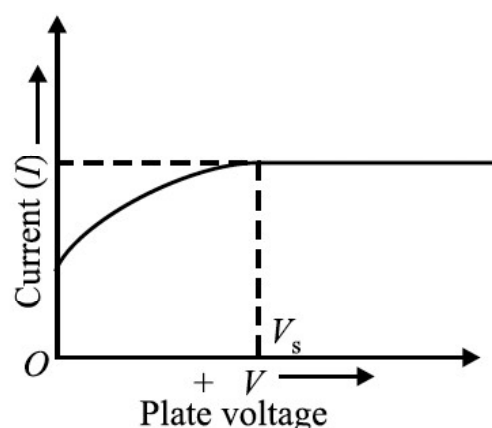


Fig. 25.2 : (a) Variation of number of photoelectrons with intensity of incident light

**Case - I : Plate A positive relative to C**

Let us first consider the case when the plate  $A$  is at a positive potential relative to  $C$ . When light of high frequency is incident on the emitter, it starts emitting electrons. Since  $A$  is at a higher potential relative to  $C$ , the emitted electrons experience an attractive force. When we increase the voltage on  $A$ , the kinetic energy of the photoelectrons increases. The current in the outer circuit shown by the microammeter depends on the number of electrons reaching the plate  $A$ . If we keep on increasing the voltage, a stage comes when all the emitted electrons are collected by the plate. The current is said to have *saturated* at this stage. If the voltage on the plate is increased further, the current remains constant in magnitude. This behaviour of current with respect to plate voltage is shown in Fig. 25.2 (b). The voltage  $V_s$  is called *saturation voltage*.



**Fig. 25.2 : (b)** Dependence of photoelectric current on plate voltage.

If the voltage on the plate is increased further, the current remains constant in magnitude. This behaviour of current with respect to plate voltage is shown in Fig. 25.2 (b). The voltage  $V_s$  is called *saturation voltage*.

**Case-II : Plate A negative relative to C**

If  $C$  is at a positive potential relative to the plate  $A$  and light of a proper frequency is incident on the emitter, photoelectrons emitted by  $C$  will experience retarding potential, which impedes their motion towards  $A$ . Some of the electrons emitted from  $C$  may still reach the plate. This gives rise to current, which is registered by the microammeter. What does this mean? If the p.d between  $A$  and  $C$  only provides the force which makes the electrons move towards the plate, then none of the electrons should have reached the plate. Since such electrons have overcome the retarding potential while moving against it to reach the plate, they have some initial kinetic energy. This is also confirmed by observed results. For any incident light of particular frequency, if the retarding potential is gradually increased in magnitude, a stage is ultimately reached when none of the electrons reach the plate and the current becomes zero.

*The minimum retarding potential for which the photoelectric current becomes zero for a particular frequency of incident light is called the stopping potential,  $V_0$  for that frequency.*

The work done by an electron  $W$  against the stopping potential  $V_0$  is  $eV_0$  where  $e$  is electronic charge. This work is done by the electron at the expense of its kinetic energy. So, we can write

$$eV_0 = \frac{1}{2}mv_{\max}^2$$

The stopping potential  $V_0$ , was found by Millikan to depend on the frequency of the incident light. A plot of the stopping potential ( $V_0$ ) versus the frequency of the incident light



( $\nu$ ) is shown in Fig. 25.3. You will note that there is a minimum cutoff frequency  $\nu_0$  below which ejection of electrons is not possible. It is called *threshold frequency*.

$\nu_0$  is the threshold frequency for metal A. Therefore electron emission for metal A takes place for frequencies greater than threshold frequency ( $\nu > \nu_0$ ). The threshold frequencies are different for different metals. The  $\nu'_0$  is the threshold frequency of metal B. So photoelectric effect takes place for the frequencies greater than threshold frequency ( $\nu > \nu'_0$ ) for metal B.

To study the effect of frequency of incident light on stopping potential, Millikan adjusted the intensity of light at a fixed value for various frequencies and studied the variation of photoelectric current with anode potential. He obtained different values of stopping potential for different frequencies of incident light. Moreover,

the stopping potential is more negative for higher frequencies, as shown in Fig. 25.4. This implies that if the frequency of the incident light increases, the maximum kinetic energy of the photoelectrons also increases. Therefore, with increasing frequency, greater retarding potential is required to completely stop the movement of photoelectrons towards the anode.

From Fig. 25.4, we can observe that the stopping potentials are in the order  $V_{03} > V_{02} > V_{01}$  for the frequencies of the order  $\nu_3 > \nu_2 > \nu_1$ . The greater the frequency of incident light, greater the kinetic energy of photoelectrons, hence more stopping potential is required to stop them completely.

This experiment also established that there exists a minimum cut-off frequency  $\nu_0$  for which stopping potential is zero. Moreover, photo emission begins as soon as light is incident on the material, i.e. photo emission is instantaneous, even if the incident light is dim. Now it is known that time lag between incident light and emission of photoelectrons from the emitter is of the order of  $10^{-9}$  s.

These observations can be summarised as follows :

- *The maximum velocity of photoelectrons increases with frequency of incident light and depends on the nature of emitter material.*
- *The maximum velocity of photoelectrons does not depend on the intensity of incident light.*

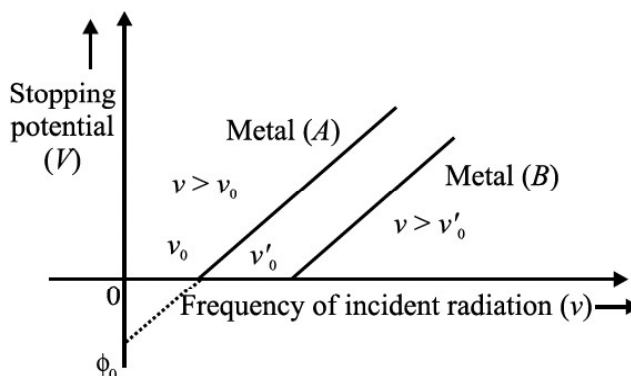


Fig. 25.3 : Stopping potential versus frequency of incident light.

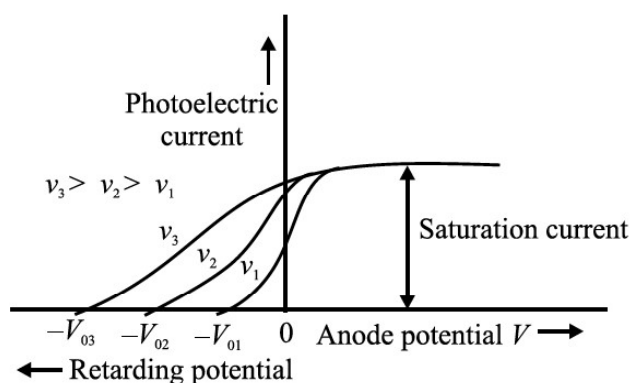


Fig. 25.4 : Photo electric current

- For every material, there exists a threshold frequency below which no photoelectrons are emitted.
- For a particular frequency, the number of photoelectrons emitted per unit area of the emitting surface is proportional to the intensity of the incident light.
- There is practically no time lapse ( $\sim 10^{-9}$  s) between the incidence of light on the metal and emission of electrons from it. In other words, photoelectric emission is an instantaneous process.

### Intext Questions 25.1

1. State whether the following statements are true or false :
  - (a) In thermionic emission, electrons gain energy from photons.
  - (b) The maximum velocity of photoelectron is independent of the frequency of incident radiation.
  - (c) There exists a frequency  $\nu_0$  below which no photoelectric effect takes place.
2. Refer to Fig. 25.3 and interpret the intercepts on  $x$  and  $y$ -axes and calculate the slope.
3. Draw a graph showing the variation of stopping potential ( $-V_0$ ) with the intensity of incident light.

## 25.2 EINSTEIN'S THEORY OF PHOTOELECTRIC EMISSION

In 1905, Einstein proposed a simple but revolutionary explanation for the photoelectric effect. He assumed that light consists of bundles of energy, called photons and viewed photoelectric effect as a collision between a photon and a bound electron.

The energy  $E$  of a single photon is given by

$$E = h\nu \quad (25.2)$$

where  $\nu$  is the frequency of the incident light and  $h$  is Planck's constant. Let us now assume that a photon of energy  $h\nu$  is incident on the metal surface. Suppose  $\phi_0$  is the energy needed for an electron to come out of the metal surface. As you have studied earlier, this energy is also called the *work function* of the conductor. *The work function of a conductor is the minimum energy required by an electron to come out of the conductor surface.*

The typical values of work function for a few metals are given (in eV.) in Table-25.1, along with the corresponding threshold frequency ( $\nu_0$ ).

**Table - 25.1: Work function and threshold frequencies of some typical metals**

Metal	$\phi_0$ (eV)	$\nu$ (Hz)
Sodium	2.5	$6.07 \times 10^{14}$
Potassium	2.3	$5.58 \times 10^{14}$
Zinc	3.4	$8.25 \times 10^{14}$
Iron	4.8	$11.65 \times 10^{14}$
Nickel	5.9	$14.32 \times 10^{14}$

### Robert A. Millikan (1868-1953)

Robert Andrews Millikan was born on March 22, 1868 in U.S.A. During his undergraduate course, his favourite subjects were Greek and Mathematics. But after his graduation in 1891, he took, for two years, a teaching post in elementary physics. In this period, he developed interest in the subject. He received his Ph.D. (1895) for research on polarization of light emitted by incandescent surfaces.



Millikan spent a year (1895-1896) in Germany, at the Universities of Berlin and Göttingen. He returned at the invitation of A.A. Michelson to take appointment as his assistant at the newly established Ryerson Laboratory at the University of Chicago (1896). He became Professor at that University in 1910, a post which he retained till 1921. As a scientist, Millikan made numerous momentous discoveries in the fields of electricity, optics, and molecular physics. His earliest major success was the accurate determination of the charge carried by an electron, using the elegant “falling-drop method”. He also proved that this quantity was a constant for all electrons demonstrating the quantised nature of charge.

He also verified experimentally Einstein’s photoelectric equation, and made the first direct photoelectric determination of Planck’s constant  $h$ . Throughout his life, Millikan remained a prolific author, making numerous contributions to scientific journals. He was awarded the Nobel Prize in Physics in 1923.

What do you think would happen when a photon of energy  $E$  ( $> \phi_0$ ) strikes the metal surface? We expect that out of the total energy  $E$ , an amount  $\phi_0$  would be used up by the electron to come out of the metal surface. The difference in energy, i.e.  $(E - \phi_0)$ , =  $K_{\max}$  would then be imparted to the emitted electron in the form of kinetic energy. (The electron may lose some energy in internal collisions before it escapes from the metal surface.) Mathematically, we can write

$$h\nu = \phi_0 + K_{\max} \quad (25.3)$$

### Albert Einstein (1879-1955)

Albert Einstein was born in Wurttemberg, Germany, on March 14, 1879. In 1901, he acquired Swiss citizenship and, as he was unable to find a teaching post, he accepted a position as technical assistant in the Swiss Patent Office. During his stay at the Patent office, in his spare time, he produced much of his remarkable work, including the theory of photoelectric effect and the special theory of relativity.



In 1909 he became Professor Extraordinary at Zurich. In 1911 he accepted the post of Professor of Theoretical Physics at Prague but returned to Zurich in the following year to fill a similar post. In 1914, he was appointed Director of the Kaiser Wilhelm Physical Institute and Professor in the University of Berlin. He became a German citizen in 1914. He was awarded the Nobel Prize in Physics in 1921 for his theory of photoelectric effect, though he is more famous for his theory of relativity. He remained in Berlin until 1933, when he renounced his citizenship for political reasons and immigrated to take the position of Professor of Theoretical Physics at Princeton in USA.

He became a US citizen in 1940 and retired from his post in 1945. He spent the later years of his life working on General Theory of Relativity and Unification of basic Forces. Einstein nurtured scientific humanism. He protested to President Roosevelt against the use of nuclear bombs for destruction of humanity. He is considered the greatest scientist to have ever walked on this planet and named scientist of the millennium.

Let us now see how observed results can be explained on the basis of this theory.

Let us take

$$\phi_0 = h\nu_0$$

Then Eqn. (25.3) takes the form

$$K_{\max} = \frac{1}{2}mv^2 = h(\nu - \nu_0) \quad (25.4)$$

This equation implies that

- For  $K_{\max}$  to be positive,  $\nu$  must be greater than  $\nu_0$ . That is, the incident light must have frequency above the threshold frequency.
- $K_{\max}$  is linearly proportional to  $(\nu - \nu_0)$
- An increase in the intensity of incident light of frequency  $\nu$  corresponds to an increase in the number of photons. Each and every photon has same energy; there is no increase in the energy of photoelectrons. However, the no. of emitted electrons and hence photocurrent will increase with increase in intensity.
- Since photoelectric effect is produced by collisions between photons and electrons, the energy transfer from photons is instantaneous, i.e. there is almost no time lag.

- Since work function is a characteristic property of a material,  $\nu_0$  is independent of the intensity of incident light.

We see that Einstein's theory of the photoelectric effect successfully explains its physical origin.

To understand these concepts and get an idea about the values of physical parameters, go through the following examples carefully.

### Example 25.1

Sodium has a work function of 2.3 eV. Calculate (i) its threshold frequency, (ii) the maximum velocity of photoelectrons produced when sodium is illuminated by light of wavelength  $5 \times 10^{-7}$  m, (iii) the stopping potential for light of this wavelength. Given  $h = 6.6 \times 10^{-34}$  J s,  $c = 3 \times 10^8$  m s<sup>-1</sup>, 1 eV =  $1.6 \times 10^{-19}$  J, and mass of electron  $m = 9.1 \times 10^{-31}$  kg.

**Solution :**

- (i) The threshold frequency is given as  $h\nu_0 = \phi_0$ . Here,  $h = 6.6 \times 10^{-34}$  Js and  $\phi_0 = 2.3$  eV =  $2.3 \times 1.6 \times 10^{-19}$  J.

$$\therefore \nu_0 = \frac{\phi_0}{h} = \frac{2.3 \times 1.6 \times 10^{-19} \text{ J}}{6.6 \times 10^{-34} \text{ Js}} = 5.6 \times 10^{14} \text{ Hz}$$

- (ii) From Einstein's photoelectric equation, we know that

$$h\nu = \phi_0 + K_{\max} = \phi_0 + \frac{1}{2}mv_{\max}^2$$

Since  $\nu = \frac{c}{\lambda}$ , we can write

$$E = h \times \frac{c}{\lambda} = \phi_0 + \frac{1}{2}(mv_{\max}^2)$$

where  $c$  is velocity of light and  $\lambda$  is its wavelength. On substituting the given values, we get

$$\therefore E = \frac{(6.6 \times 10^{-34} \text{ Js})(3 \times 10^8 \text{ ms}^{-1})}{5 \times 10^{-7} \text{ m}} = 3.96 \times 10^{-19} \text{ J}$$

$$\Rightarrow 3.96 \times 10^{-19} = 2.3 \times 1.6 \times 10^{-19} + \frac{1}{2}mv_{\max}^2$$

$$= 3.68 \times 10^{-19} + \frac{1}{2}mv_{\max}^2$$

$$v_{\max}^2 = \frac{2 \times 0.28 \times 10^{-19}}{m} = \frac{2 \times 0.28 \times 10^{-19}}{9.1 \times 10^{-31}}$$

$$v_{\max} = \sqrt{\frac{0.56 \times 10^{-19} \text{ J}}{9.1 \times 10^{-31} \text{ kg}}} = 2.5 \times 10^5 \text{ ms}^{-1}$$

(iii) The stopping potential  $V_0$  is given as

$$eV_0 = \frac{1}{2}mv_{\max}^2$$

$$V_0 = \frac{0.28 \times 10^{-19} \text{ J}}{1.6 \times 10^{-19} \text{ J V}^{-1}} = 0.18 \text{ V}$$

You may now like to answer some simple questions

### Intext Questions 25.2

1. Calculate the momentum of a photon of frequency  $\nu$ .
2. If the wavelength of an electromagnetic radiation is doubled, how will be the energy of the photons change?
3. The intensity of incident radiation is doubled. How will it affect the kinetic energy of emitted photoelectrons.

In the previous section, you have studied Einstein's theory of photoelectric effect and learnt that light consists of photons. You have also learnt that the phenomena of interference and diffraction can be explained on the basis of wave theory of light. This duality in the nature of light came to be accepted by the physicists in the early 20<sup>th</sup> century. Thinking about the wave-particle duality of light, de Broglie asked himself the question : If light exhibits dual nature, will particles of matter also not act like waves? Successful resolution of this question led to de Broglie hypothesis.

### 25.3 THE DE BROGLIE HYPOTHESIS

As a young graduate student, de Broglie argued with a great amount of insight that since nature loves symmetry and simplicity in physical phenomena, ordinary "particles" such as electrons, and protons should also exhibit wave characteristics under certain circumstances. His argument runs as follows : Light is an electromagnetic radiation and exhibits wave-particle duality. Therefore, Einstein's mass-energy equivalence relation ( $E = mc^2$ ), which essentially treats light as quantum of photon, a particle, can hold only if matter also exhibits wave character. He therefore proposed that the wavelength and frequency of matter waves should be determined by the momentum and energy of the particle in exactly the same way as for photons :  $E = pc$  and the associated wavelength  $\lambda$  of a particle having momentum  $p$  is given by

$$\lambda = \frac{h}{p} \quad (25.5)$$

Since the momentum of such a particle is given by  $p = mv$ , we can write

$$\lambda = \frac{h}{mv} \quad (25.6)$$

$\lambda$  is called deBroglie wavelength. Eqn. (25.5) is a complete statement of waveparticle duality. It implies that a particle with a momentum  $p$  can exhibit wavelike properties and the wavelength of the associated matter waves is  $h/p$ . The converse is also true, i.e., a wave of wavelength  $\lambda$  can exhibit particle-like properties and the momentum of the wave-matter is  $h/\lambda$ .

This hypothesis, submitted as Ph.D. Thesis was initially rejected by the examiners. However, soon, experimental evidence proved de Broglie's argument. This has a very important inspirational lesson for us : We must keenly analyse every statement and try to seek experimental evidence.

The actual wavelength of anything macroscopic is incomprehensively small, as you can see by calculating it for a cricket ball. The case is quite different for elementary particles such as the electron. An electron has energy  $E$  when accelerated through potential difference  $V$ . Hence, we can write

$$\frac{1}{2}mv^2 = qV \quad (25.7)$$

or

$$v = \sqrt{\frac{2qV}{m}}$$

so that

$$mv = p = \sqrt{2qmV} \quad (25.8)$$

On combining this result with Eqn. (25.5), we find that de-Broglie wavelength is given by

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2qmV}} \quad (25.9)$$

The constants appearing in Eqn. (25.9) have the values:  $h = 6.625 \times 10^{-34}$  Js,  $q = 1.602 \times 10^{-19}$  C and  $m = 9.11 \times 10^{-31}$  kg. On substituting these values in Eqn. (25.9), we obtain

$$\lambda = \frac{6.625 \times 10^{-34} \text{ Js}}{\sqrt{2 \times (1.602 \times 10^{-19} \text{ C}) \times (9.11 \times 10^{-31} \text{ kg}) \times \sqrt{V}}}$$

$$\begin{aligned}
 &= \frac{12.3}{\sqrt{V}} \times 10^{-10} \text{ m} \\
 &= \frac{12.3}{\sqrt{V}} \text{ \AA} \qquad (25.10)
 \end{aligned}$$

It means that if an electron is accelerated through a potential difference of 100V, its wavelength will be given by

$$\begin{aligned}
 \lambda &= \frac{12.3}{\sqrt{100}} \text{ \AA} \\
 &= 12.3 \text{ \AA}
 \end{aligned}$$

This is also the wavelength of an electron of energy 100 eV. You can easily verify this using the relation

$$\lambda = \frac{h}{(2meE)^{1/2}}$$

The wavelength of matter waves associated with 100 eV electrons lies in the X-ray region and is of the same order as the inter atomic separation in a solid. We therefore expect these to undergo diffraction by a crystal lattice.

The first experimental evidence of matter waves came from the work of Davisson and Germer, who were studying scattering of electrons by crystals. Let us learn about it now.

### Louis Victor de Broglie (1892-1987)

Louis de Broglie was born at Dieppe, France on 15<sup>th</sup> August, 1892. He first studied the arts and took his degree in history in 1910. Then, as his liking for science prevailed, he studied for a science degree, which he gained in 1913. In 1924 at the Faculty of Sciences at Paris University, he submitted a thesis *Recherches sur la Theorie des Quanta* (Researches on the quantum theory), which gained him his doctor's degree. This thesis contained a series of important findings, which he had obtained in the course of about two years. The ideas set out in that work served the basis for developing *wave mechanics*, a theory which has transformed our knowledge of physical phenomena on the atomic scale.



In 1929 he was awarded the Nobel Prize for Physics for his discovery of the wave nature of electrons.



### 25.3.1 Experimental Evidence for Existence of de Broglie Waves - Davisson-Germer Experiment

The schematic diagram of Davisson-Germer experiment is shown in Fig. 25.5. The set up consists of a filament  $F$ , which serves as a source of electrons. The electrons emitted from this filament are made to pass through a set of metal diaphragms  $D$  having a number of slits. The electrons emitted by the filament come out in various directions. The metal diaphragms serve to collimate these electrons. Only those electrons which are able to pass through the slits in the various diaphragms are able to come out.

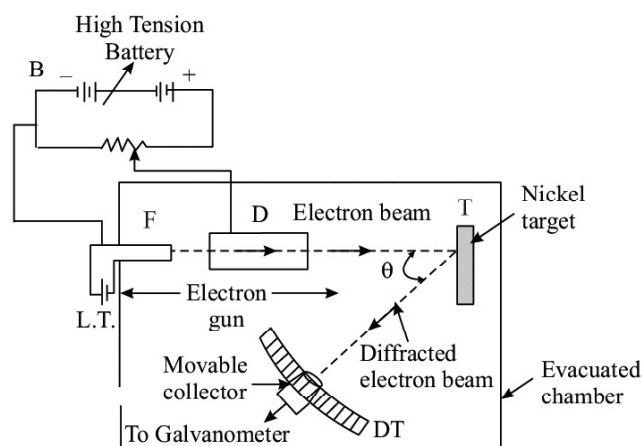


Fig. 25.5 : Experimental set up to verify the existence of matter waves

Note that the energy of the collimated stream of electrons is controlled by changing the magnitude of the accelerating voltage. The beam of electrons is made to fall perpendicularly on a single crystal of nickel  $T$ . The set-up also contains a detector  $DT$  which can be placed at any particular angle  $\theta$  with respect to the normal to the target crystal. This detector determines the intensity of the reflected beam. Note that there is nothing special in the choice of nickel.

The intensity of the scattered electron beam is measured for different values of angle of scattering  $\theta$ , which is the angle between the incident and the scattered electron beams.

The Fig. 25.6 shows the variation of detector current versus kinetic energy of incident electrons. The experiment was performed by varying kinetic energy of electrons. It was observed that a strong peak appeared in the intensity of scattered electron for kinetic energy of electrons of 54 eV at a scattering angle  $50^\circ$ . If you calculate the wavelength of these electrons, you will get

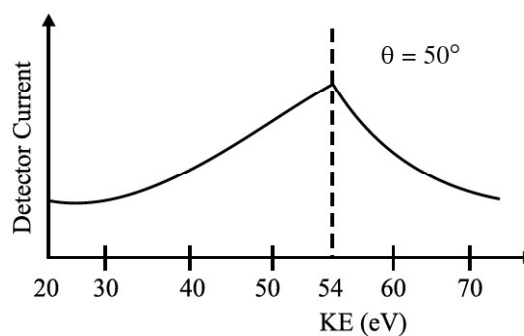


Fig. 25.6 : Plot of detector current versus kinetic energy of electrons

$$\lambda = \frac{6.62 \times 10^{-34} \text{ Js}}{\left[ 2 \times (9.1 \times 10^{-31} \text{ kg}) \times 54 \times 1.6 \times 10^{-19} \text{ J} \right]^{1/2}}$$

$$= 1.67 \text{ \AA}$$

This shows that an excellent agreement between the theoretical value and experimentally obtained value of de Broglie wavelength.

### Story of Davisson and Germer's experiment

Germer had recorded in his notebook that he discovered a crack in the vacuum trap in the electron scattering apparatus on Feb. 5, 1925 when he was working with Clinton Davisson at Western Electric, New York, U.S.A. This was not the first time their equipment had broken, and not the first time they had “resurrected” their precious nickel crystal by **heating it in vacuum and hydrogen.**



This particular break and the subsequent method of repair, however, had a crucial role to play in the later discovery of electron diffraction. By 6 April 1925, the repairs had been completed and the tube put back into operation. During the following weeks, as the tube was run through the usual series of tests, results very similar to those obtained four years earlier were obtained. Then suddenly, in the middle of May, unprecedented results began to appear. This puzzled Davisson and Germer so much that they halted the experiments a few days later, cut open the tube and examined the target (with microscopist F. F. Lucas) to see if they could detect the cause of new observations. What they found was this: the polycrystalline form of nickel target had been changed by the extreme heating until it had formed about ten crystal facets in the area from which the incident electron beam was scattered. Davisson and Germer surmised that the new scattering pattern must have been caused by the new crystal arrangement of the target. In other words, they concluded that it was the arrangement of the atoms in the crystals, not the structure of the atoms that was responsible for the new intensity pattern of scattered electrons.

During the summer of 1926, Davisson and his wife had planned a vacation trip to relax and visit relatives in England. Something was to happen on this particular trip. Theoretical physics was undergoing fundamental changes at this time. In the early months of 1926, Erwin Schrodinger's remarkable series of papers on wave mechanics appeared, following Louis de Broglie's papers of 1923-24 and Albert Einstein's quantum gas paper of 1925. These papers were the subject of lively discussions at the Oxford meeting of the British Association for the Advancement of Science.

Davisson, who generally kept abreast of recent developments in his field but appears to have been largely unaware of these developments in quantum mechanics, attended this meeting. He was surprised when he heard a lecture by Born in which his own and Kunsman's (platinum target) curves of 1923 were cited as confirmatory evidence for de Broglie's electron waves!

Davisson shared the 1937 Nobel Prize for Physics with G.P. Thomson (son of J.J. Thomson).

**Example 25.2**

An electron is accelerated through a potential difference of 182 V. Calculate its associated wavelength.

**Solutions :**

We know that deBroglie wavelength,  $\lambda = \frac{h}{p} = \frac{12.3}{\sqrt{V}} \text{ \AA}$ . Here  $V = 182 \text{ V}$

$$\therefore \lambda = \frac{12.3}{\sqrt{182}} \text{ \AA} = \frac{12.3}{13.5} = 0.91 \text{ \AA}$$

**Example 25.3**

Calculate the maximum kinetic energy of the emitted photoelectrons when light of frequency  $\nu = 10^{15} \text{ Hz}$  is incident on a zinc plate. The work function of zinc is 3.4 eV.

**Solution :**

From Einstein's relation, we recall that

$$h\nu = \phi_0 + K_{\max}$$

For this problem,  $h = 6.625 \times 10^{-34} \text{ Js}$ ,  $\nu = 10^{15} \text{ Hz}$ ,  $E = h\nu = 6.625 \times 10^{-34} \times 10^{15}$   
 $= 6.625 \times 10^{-19} \text{ J}$  and  $\phi_0 = 3.4 \text{ eV} = 3.4 \times 1.602 \times 10^{-19} \text{ J} = 5.4468 \times 10^{-19} \text{ J}$

$$\therefore K_{\max} = E - \phi_0 = (6.625 - 5.447) \times 10^{-19} \text{ J} = 1.178 \times 10^{-19} \text{ J}$$

**Intext Questions 25.3**

- State whether the following statements are true or false:
  - According to deBroglie, stationary particles exhibit wave-like characteristics.
  - Matter waves are the same thing as deBroglie waves.
  - Very poor resolution can be obtained in a microscope using energetic electrons by lowering deBroglie wavelengths associated with electrons.
- A 50 g ball rolls along a table with a speed of  $20 \text{ cm s}^{-1}$ . How large is its associated wavelength? Given  $h = 6.625 \times 10^{-34} \text{ Js}$ .
- Why can we not observe de Broglie wavelength associated with a cricket ball?

**25.4 PARTICLE NATURE OF LIGHT! THE PHOTON**

It is evident from the photoelectric effect that interaction of radiation with matter behaves as if it is made up of quanta or packets, each of energy  $h\nu$ , where  $h$  is plank constant and  $\nu$  is the frequency of radiation. This quanta or particles further named as photon. [So light

consists of small tiny particles called photons. This confirms the particle nature of light].

These photons also associated with momentum  $\frac{h\nu}{c}$ , where  $c$  is speed of the light. A definite value of energy as well as momentum is a strong sign that light quantum can be associated with a particle. The particle like behaviour also confirmed by Compton on scattering of x-rays from electrons.

The following points can be drawn from the photon picture of electromagnetic radiation.

1. The light interaction with matter behaves as if it is made up of particles called photon.
2. Each photon having the energy  $h\nu$ , where  $h$  is plank constant and  $\nu$  is the frequency of light and momentum  $P = \frac{h\nu}{c}$ ,  $c$  is the speed of light.
3. The photon energy depends on the frequency and independent of intensity of radiation. By increasing the intensity, there is only increase in number of photons having same energy.
4. Photons are electrically neutral and do not deflected by electric and magnetic fields.
5. In a photon-particle (electron) collision, the total energy and momentum are conserved. But number of photons are not conserved. The photons may be absorbed or new photons may be created.

## WHAT YOU HAVE LEARNT

- Emission of electrons from a metal when light of proper frequency incident on its surface is called photoelectric emission.
- In photoelectric emission, electrons gain energy from light.
- The stopping potential increases with increase in frequency of incident light.
- There exists a frequency  $\nu_0$  for every material below which no photoelectric effect takes place.
- The maximum velocity of photoelectrons increases with increasing frequency of incident light but is independent of the intensity of incident light.
- The number of photoelectrons emitted from each square centimeter of the emitting surface for any particular frequency is proportional to the intensity of incident light.
- Einstein assumed light to consist of photons, each having energy  $h\nu$ , where  $\nu$  is frequency and  $h$  is Planck's constant.
- The saturation current of a phototube increases with increasing intensity of the incident light.
- Particles in motion have waves associated with them. The wavelength is given by  $h/p$ , where,  $p$  is the momentum.

**TERMINAL EXERCISE**

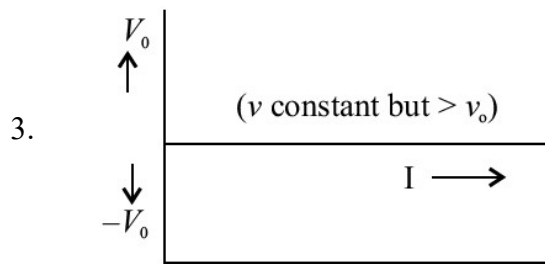
- In photoelectric emission, what happens to the incident photons?
- What is the difference between a photon and a matter particle?
- Why is the wave nature of matter not apparent in daily life?
- How is velocity of photoelectrons affected if the wavelength of incident light is increased?
- The threshold frequency of a metal is  $5 \times 10^{14}$  Hz. Can a photon of wavelength  $6000 \text{ \AA}$  emit an energetic photoelectron?
- Does the threshold frequency for a metal depend on the incident radiations?
- What are the various uses of photocell?
- What was the aim of Davisson and Germer's experiment? On what principle does it depend?
- Describe the experiment used for studying the photoelectric effect.
- Explain the terms (a) Saturation voltage and (b) Stopping potential.
- State the laws of photoelectric emission.
- Describe the salient features of Einstein's theory of photoelectric effect.
- Explain Einstein's relation:  $h\nu = E_0 + K_{\max}$
- Calculate the wavelength associated with electrons moving with a velocity  $v = 1 \times 10^8 \text{ ms}^{-1}$ . Take mass of electron =  $9.1 \times 10^{-31} \text{ kg}$  and  $h = 6.6 \times 10^{-34} \text{ J.s}$ .
- Describe an experiment which verifies the existence of deBroglie waves.
- Show that the deBroglie wavelength associated with electrons accelerated through a potential  $V$  is given by the relation;

$$\lambda = \frac{12.3}{\sqrt{V}} \text{ \AA}$$

**ANSWERS TO INTEXT QUESTIONS****25.1**

- (a) False (b) False (c) True
- $x$  – intercept gives the threshold frequency  
 $y$  – intercept gives  $e \times$  work function ( $\phi_0$ )

$$V_0 = \frac{h}{e} \nu - \frac{h}{e} \nu_0, \text{ slope of graph gives } \frac{h}{e}$$


**25.2**

1.  $\lambda = \frac{h}{p} \Rightarrow p = \frac{h}{\lambda} = \frac{h}{c/v} = \frac{hv}{c}$

2.  $E = hc / \lambda$

If  $\lambda$  is doubled, E will become half

3. It is unchanged.

**25.3**

1. (a) false (b) True (c) True

2.  $P = mv$  and  $\lambda = \frac{h}{p}$

Here  $m = 50 \text{ g} = 0.05 \text{ kg}$  and  $v = 20 \text{ cms}^{-1}$

$\therefore \lambda = 6.6 \times 10^{-32} \text{ m}$

3. From Eqn. (25.6) it is clear that if mass  $m$  is large, the value of  $\lambda$  will be small. Same is the case with cricket ball.

<b>ANSWERS TO TERMINAL EXERCISE</b>
-------------------------------------

14.  $7.25 \text{ \AA}$



# STRUCTURE OF THE ATOM

## INTRODUCTION

We have studied about mechanical, thermal, electrical and magnetic properties of matter. Have you ever thought, why do some metals show conducting current when light falls on them? To understand such properties of materials, we recall that atoms are building blocks of all forms of matter. Matter has definite structure on microscopic level which is beyond the reach of our sense of seeing. Our understanding of the structure of atom has evolved over a period of time. In this chapter, we will study different atomic models right from Rutherford's model, Bohr's atomic model and atomic spectrum of hydrogen atom.

### OBJECTIVES

After studying this lesson, you should be able to understand the

- *Rutherford's scattering experiment and its findings;*
- *Bohr's first orbit radius and velocity of an electron in it;*
- *mathematical expression for the energy of an electron in a hydrogen atom;*
- *energy level diagram of a hydrogen atom and its spectrum.*

## 26.1) INTRODUCTION TO THE CONCEPT OF ATOM

In ancient times Democritus in Greece and Kanad Maharshi in India tried to explain the changes around us in terms of particles. But the exact theory of atom was presented by John Dalton, an English Chemist in 1808. He described atom as the smallest, indivisible particle endowed with all the properties of element and takes part in chemical reactions. Dalton's atom was an ultimate particle having no structure. This idea was accepted by the scientists in the nineteenth century as they knew nothing about the structure of atoms. The discovery of electrons by J.J. Thomson in 1897, while studying discharge of electricity through gases at low pressures, they thought that atoms have a structure and negatively charged electrons are constituents of all atoms. It was also thought that the positively charged constituent of atoms carried the entire mass. Thomson suggested the *plum pudding model* of atom (Fig. 26.1). According to this, an atom is a tiny, uniformly charged positive ball in which negatively charged electrons are suitably placed to make it neutral. It seemed perfectly reasonable at that time.

Our understanding of the structure of atom since the times of Thomson has improved considerably. Due to the pioneering

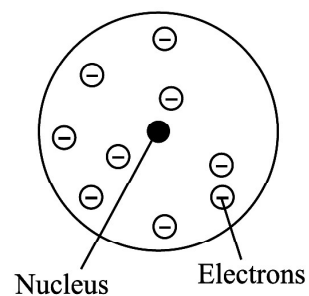


Fig. 26.1 : Plum-pudding model of atom

works of Lord Rutherford, Niels Bohr, James Chadwick, Pauli, Schrodinger and others, our concept of new world of sub-atomic particles came into existence and has led to the invention of epoch-making new technologies, like micro-electronics and nanotechnology.

### 26.1.1 Rutherford's model on scattering of $\alpha$ -particle

Lord Rutherford and his two students (Geiger and Marsden) have performed an experiment in which a beam of  $\alpha$ -particles was bombarded on a thin gold foil. The experimental arrangement is shown in Fig. 26.2.

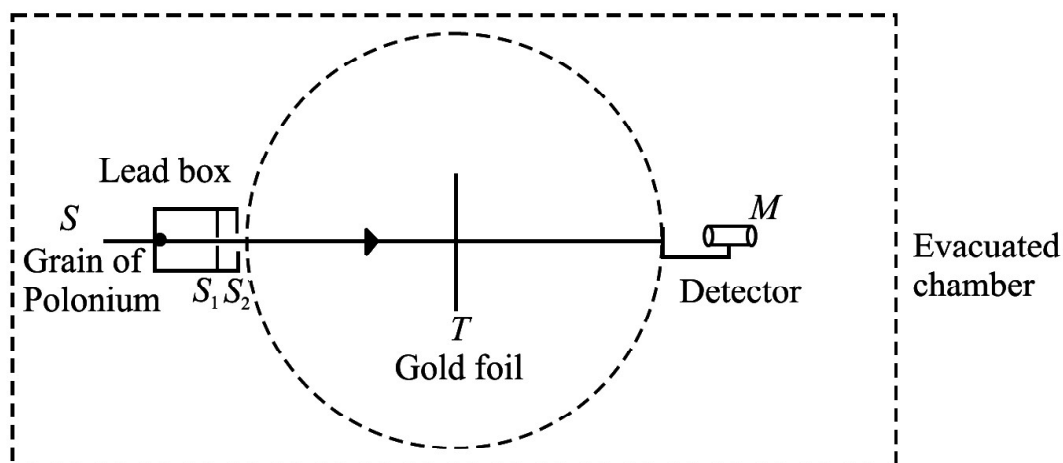


Fig. 26.2 : Schematics of experimental set up used for  $\alpha$ -particle scattering

A well collimated fine beam of  $\alpha$ -particles coming from  $S_1$  and  $S_2$  slits fall on a thin gold foil (T), from a source S. The scattered  $\alpha$ -particles were received on a ZnS fluorescent screen, which produced a visible flash of light when struck by  $\alpha$ -particles (and acted as detector), backed by a low power microscope (M). The detector was capable of rotation on a circular scale with T at the centre.

The whole apparatus was enclosed in an evacuated chamber to avoid collisions of  $\alpha$ -particles with air molecules. Geiger and Marsden observed that most of the  $\alpha$ -particles suffered only small deflections, as expected. But a few got deflected at large angles ( $90^\circ$  or more). Some of them (1 in 8000) even got deflected at  $180^\circ$ . Fig. 26.3 presents the experimental results. The large angle scattering of  $\alpha$ -particles could not be explained on the basis of Thomson model of atom.

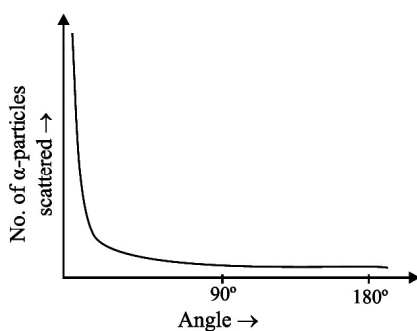


Fig. 26.3 : Experimental result of Rutherford's experiment

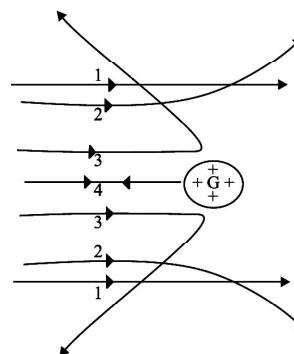


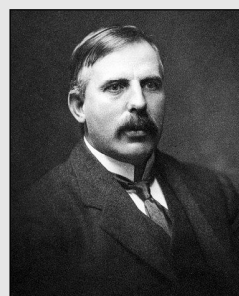
Fig. 26.4 : Paths traversed by  $\alpha$ -particles scattered by a gold foil



It was expected that if Thomson model was correct, most of the particles would go straight through the foil, with only minor deviation from the original path. To explain large angle scattering, Lord Rutherford suggested the nuclear model of atom. He argued that  $\alpha$ -particles which pass at a large distance from the nucleus experience negligible repulsive force and un-deflected. However, closer to the nucleus a  $\alpha$ -particle comes, greater force of repulsion and hence gets deflected at a greater angle. A few  $\alpha$ -particles which proceed for a head-on collision towards the nucleus are scattered back by  $180^\circ$  along its direction of approach, as indicated by  $\alpha$ -particle 4 in Fig. 26.4.

### Lord Rutherford (1871–1937)

Born in New Zealand, Rutherford studied under J.J. Thomson at the Cavendish Laboratory in England. His pioneering work on atom is a defining landmark. He developed Becquerel's discovery of radioactivity into an exact science and documented proof that the atoms of heavier elements, which had been thought to be immutable, actually disintegrate (decay) into various forms of radiation. In 1898, Rutherford discovered that two quite separate types of emissions came from radioactive atoms and he named them alpha and beta rays. Beta rays were soon shown to be high speed electrons. In 1907, he showed that the alpha particle was a helium atom stripped of its electrons. He and his assistant, Hans Geiger, developed Rutherford-Geiger detector to electrically detecting particles emitted by radioactive atoms. With this he could determine important physical constants such as Avogadro's number, the number of atoms or molecules in one gram-mole of material.



In 1911, Rutherford proposed the nuclear model of atom; that almost entire mass of an atom is concentrated in a nucleus  $10^{-5}$  times the atom itself and electrons revolve around it. This second great work won him the **Nobel Prize** in chemistry in 1908.

Eminent Indian physicist, educationist and philosopher Dr. D.S. Kothari was one of his students and worked on pressure ionisation in stars.

#### 26.1.2 Nuclear Model of Atom

Rutherford argued that large angle scattering of  $\alpha$ -particles can be explained only by stipulating the presence of a hard, positively charged core of atom. Thus he proposed a new model of atom with following characteristics:

- (i) The entire charge and most of the mass of the atom is confined in a very small ( $\sim 10^{-15}$  m) central region, called the *nucleus*.
- (ii) The negatively charged electrons revolve at a distance around it in such a way that the atom as a whole is electrically neutral and stable.

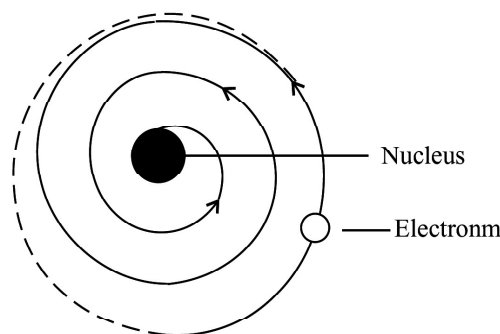
The nuclear model of atom proposed by Rutherford faced some difficulties. Some of the consequences of Rutherford's model contradicted the experimental observations.

**Stability of the atom :** We know that electrons are negatively charged. These are attracted by the nucleus and get accelerated. An accelerated charged particle, according

to classical wave theory, emits electromagnetic radiations. Hence, the revolving electrons should lose energy eventually and spiral into the nucleus as shown in the Fig. 26.5. This would have made the atom short-lived and contradicted the observed stability of matter.

**Frequency of electromagnetic radiation:**

The electron spiraling towards the nucleus will emit electromagnetic radiations of all frequencies giving rise to a continuous spectrum. But experiments show that atoms emit radiations of certain well defined frequencies only (line spectra).



**Fig. 26.5 :** Motion of electrons in Rutherford's nuclear model of atom

Based on the above discussion, one can conclude that nuclear model of atom could not explain the experimental facts. But this has contributed significantly to our understanding and was the first landmark in the right direction. To overcome the said above difficulties of Rutherford model, in 1913, Neils Bohr proposed a model of atomic structure, based on quantum ideas of Max Planck.

### Intext Questions 26.1

1. In Rutherford's scattering experiment which rays are used for target bombardment?
2. The large angle scattering of  $\alpha$ -particles indicates the presence ————— particles.
3. What experimental observations that could not be explained by Rutherford's model?

## 26.2 BOHR'S MODEL FOR THE HYDROGEN ATOM

Bohr's model describes the structure of atom and its stability. It is a highly successful model in explaining the observed spectrum of the hydrogen atom. Bohr started with the planetary model of atom. However, to overcome the problems Bohr made several assumptions. These are known as Bohr's postulates. The four postulates are:

- (i) Electrons in an atom move in circular orbits around the nucleus with the centripetal force supplied by the Coulomb force of attraction between the electron and the nucleus. Mathematically, we can write

$$\frac{mv^2}{r} = \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r^2} \quad (26.1)$$

The term  $Z$  denotes the number of positive charges in the nucleus.

- (ii) Of the infinite number of possible circular orbits, only those orbits are allowed for which the value of orbital angular momentum of the electron is an integral multiple of  $h/2\pi$

$$|\mathbf{L}| = mvr = \frac{nh}{2\pi} \quad (26.2)$$

$\mathbf{L}$  is the orbital angular momentum, equal to  $mvr$  for a circular orbit. Here  $h$  is Planck's constant and  $n$  is an integer.

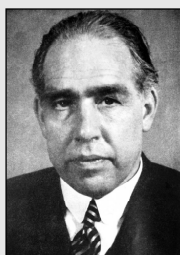
- (iii) An electron moving in an allowed orbit does not radiate any energy. In these allowed orbits, the energy of the electron is constant. These orbits are called stationary states.
- (iv) Energy is emitted by an atom only when its electron "falls" from an allowed higher energy level  $E_f$  to another allowed lower level  $E_i$ . The change in energy is the energy of the emitted photon. Similarly, an electron only absorbs radiation when it "jumps" to a higher energy  $E_f$  level from a lower energy level  $E_i$ . The change in energy of an electron can be related to the frequency ( $\nu$ ) or wavelength of the emitted or absorbed photon:

For emission and absorption process, the required equations are:

$$\Delta E = E_f - E_i = h\nu \quad (\text{For emission}) \quad (26.3 \text{ a})$$

$$\Delta E = E_i - E_f = h\nu \quad (\text{For absorption}) \quad (26.3 \text{ b})$$

### Niels Henrik David Bohr (1885-1962)



**Niels Bohr** was born in Copenhagen, Denmark. He grew up in an atmosphere most favourable to the development of his genius. His father was an eminent physiologist and was largely responsible for awakening his interest in physics while he was still at school. In the spring of 1912, he worked in Rutherford's laboratory in Manchester. He studied the structure of atoms on the basis of Rutherford's nuclear model of atom. He succeeded in working out and presenting a picture of atomic structure that explained atomic spectra of hydrogen atom.

In 1916, he was appointed Professor of Theoretical Physics at Copenhagen University, and in 1920 (until his death in 1962), he became head of the Institute for Theoretical Physics, established for him at that university.

Recognition of his work on the structure of atom came with the award of the **Nobel Prize** in Physics in 1922.

#### 26.2.1 Energy Levels

In order to calculate the energy of an electron in  $n^{\text{th}}$  orbit of radius  $r_n$ , Eqn. 26.1 can be repressed as:

$$\frac{mv_n^2}{r_n} = \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_n^2}$$

Here,  $v_n$  is speed of the electron. On multiplying both sides of this equation by  $mr_n^3$

we get

$$m^2 v_n^2 r_n^2 = \frac{1}{4\pi\epsilon_0} m Z e^2 r_n$$

On combining this result with Eqn. (26.2) and solving, finally we get

$$m^2 v_n^2 r_n^2 = n^2 \frac{h^2}{4\pi^2} = \frac{m}{4\pi\epsilon_0} Z e^2 r_n \quad (26.4)$$

On re-arranging terms, an expression for the radius of the  $n^{\text{th}}$  orbit is:

$$\begin{aligned} r_n &= 4\pi\epsilon_0 \frac{n^2 h^2}{4\pi^2 m Z e^2} \\ &= \frac{n^2 h^2 \epsilon_0}{Z e^2 m \pi} \quad n = 1, 2, 3, \dots \end{aligned} \quad (26.5)$$

From the above equation it is evident that the radius of an orbit is directly proportional to second power of the number of orbit. The relative values of the radii of permitted orbits are in the ratio 1 : 4 : 9 : 16 and so on. For hydrogen atom ( $Z = 1$ ), the radius of its inner most orbit is called **Bohr radius**. It is denoted by  $a_0$  and its magnitude is  $5.3 \times 10^{-11}$  m. In terms of  $a_0$ , the radii of other orbits are given by the relation

$$r_n = n^2 a_0$$

It shows that the spacing between consecutive orbits increase progressively. On inserting the value of  $r_n$  from Eqn. (26.5) in Eqn. (26.2), we get an expression for the speed of the electron in the  $n^{\text{th}}$  orbit:

$$\begin{aligned} v_n &= \frac{nh}{2\pi m r_n} = \frac{nh}{2\pi m} \cdot \frac{Z e^2 m \pi}{n^2 h^2 \epsilon_0} \\ &= \frac{1}{2} \frac{Z e^2}{\epsilon_0 n h} \end{aligned} \quad (26.6)$$

Since the potential energy of a negative charge (electron in this case) in bringing it from Infinity to a point at a distance  $r$  in a field of positive charge (nucleus in this case) is obtained by summing (integrating) the product of Coulomb force and distance

$$\begin{aligned} U &= -\frac{1}{4\pi\epsilon_0} \int_{r_n}^{\infty} \frac{Z e^2}{r^2} dr \\ &= -\frac{1}{4\pi\epsilon_0} \left[ \frac{Z e^2}{r} \right]_{r_n}^{\infty} \end{aligned}$$

$$= -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_n} \quad (26.7)$$

Since potential energy of the electron at infinity is zero, therefore from Eqn. (26.1), it implies,

$$\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_n} = mv_n^2$$

Hence, potential energy of an electron in  $n^{\text{th}}$  orbit and kinetic energies are:

$$U = -mv_n^2 \quad (26.8)$$

$$\text{K.E.} = \frac{1}{2}mv_n^2 \quad (26.9)$$

The total energy of the electron in  $n^{\text{th}}$  orbit is:

$$E = \text{K.E.} + U$$

$$= \frac{1}{2}mv_n^2 - mv_n^2$$

$$= -\frac{1}{2}mv_n^2$$

Combining the results with Eqn. 26.6 then,

$$\begin{aligned} E &= -\frac{m}{2} \left( \frac{2\pi Ze^2}{4\pi\epsilon_0 nh} \right)^2 \\ &= -\frac{m}{8\epsilon_0^2} \frac{Z^2 e^4}{n^2 h^2} \end{aligned} \quad (26.10)$$

$$= \frac{RZ^2}{n^2}; n=1, 2, 3... \quad (26.11)$$

Where  $R = \frac{me^4}{8\epsilon_0^2 h^2} \quad (26.12)$

Here R is known as Rydberg's constant.

From Eqn. 26.11, we note that:

1. Energy of an electron in various allowed orbits is inversely proportional to the square of the number of orbit.

2. The energy in an orbit is negative, which implies that the electron is bound to the nucleus.

By substituting the standard values of  $m = 9.11 \times 10^{-31} \text{ kg}$ ,  $e = 1.6 \times 10^{-19} \text{ C}$ ,  $\epsilon_0 = 0.85 \times 10^{-11} \text{ C}^2 \text{ N}^{-1} \text{ m}^{-2}$ , and  $h = 6.62 \times 10^{-34} \text{ Js}$  in Eqn. (26.11), we obtain  $R = 2.17 \times 10^{-18} \text{ J} = 13.6 \text{ eV}$ , since  $1 \text{ eV} = 1.6 \times 10^{-19} \text{ J}$ .

On using this result in Eqn. (26.11), we find that the energy of an electron in  $n^{\text{th}}$  orbit of Hydrogen atom (in eV) is given by the following equation:

$$E_n = -\frac{13.6}{n^2} \text{ eV}$$

Thus every orbit can be specified with a definite energy; the energy of the first orbit being the lowest:

$$E_1 = -13.6 \text{ eV}$$

And the highest energy state

$$E_n = 0$$

It means that different orbits represent different energy levels  $-13.6 \text{ eV}$  to  $0$ . This is depicted in Fig. 26.6.  $E = 0$  signifies that the electron is free.

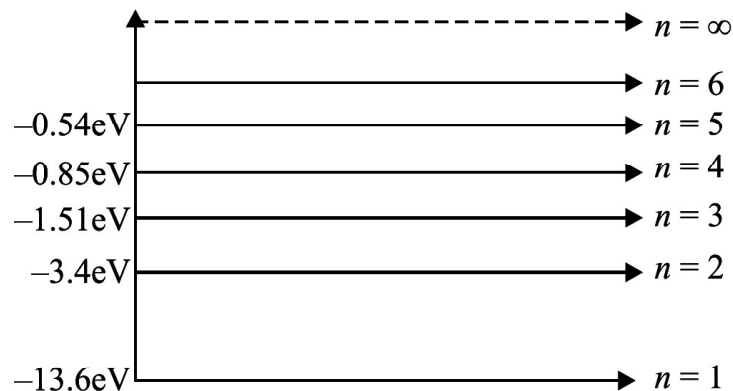


Fig. 26.6 : Energy levels in Hydrogen atom

According to Bohr's fourth postulate, the Frequency  $\nu_{mn}$  of the emitted (absorbed) radiation when the electron falls (jumps) from the  $n^{\text{th}}$  state to the  $m^{\text{th}}$  state is given by

$$\nu_{mn} = \frac{RZ^2}{h} \left( \frac{1}{m^2} - \frac{1}{n^2} \right) \quad (26.14)$$

### Fraunhofer Lines

The spectrum of sunlight, when examined carefully by a high power spectroscope, is found to be crossed by a large number of dark lines spread over the length of the continuous spectrum. Wollaston observed these lines in the year 1802. But their existence was studied by Fraunhofer on the basis of Kirchoff's laws and

named these as Fraunhofer lines. The main body of the sun emits continuous spectrum but the atmosphere of comparatively much cooler vapors and gases in the Sun's atmosphere, called the chromosphere ( $\sim 6000^\circ\text{C}$ ), absorb light corresponding to certain wavelengths. These appear as dark lines in the continuous spectrum of the sun.

Kirchhoff compared the absorbed wavelengths with the wavelengths emitted by various elements present on the earth and identified 60 terrestrial (existing on earth) elements present in the outer atmosphere of sun, e.g. oxygen, hydrogen, sodium, iron, calcium etc.

### Intext Questions 26.2

1. Which of Bohr's postulates supports classical physics and which support the ideas of quantum physics?
2. According to Bohr's theory, why did an atom did not collapse while its electrons revolved around the nucleus?
3. According to Bohr, what is happening in the atom when a photon of light if (i) emitted (ii) absorbed?
4. Write the energy of the first three orbits of hydrogen atom on the basis of Bohr's model.
5. An atom is excited to an energy level  $E_1$  from its ground state energy level  $E_0$ . What will be the wavelength of the radiation absorbed?
6. The radius of electron in the  $n^{\text{th}}$  orbit is proportional to (for hydrogen atom) \_\_\_\_\_
7. Total energy of  $n^{\text{th}}$  orbit electron of hydrogen is proportional to \_\_\_\_\_

### 26.3 HYDROGEN ATOM

Fig. 26.7 shows frequency spectrum of hydrogen atom. From the figure, it is noted that the line spectrum of hydrogen consists of many lines in different regions of the spectrum. The various lines in a particular region of spectrum are found to have a pattern and may be represented by a common formula. They are said to form a series. Let us consider the series of hydrogen spectrum.

**Lyman series** was discovered in 1906. According to Bohr, this series arises when an electron jumps to the first orbit ( $m = 1$ ) from an higher orbit ( $n = 2, 3, 4\dots$ ). The frequencies of various spectral lines of this series are given by

$$\nu_{\text{ln}} = \frac{R}{h} \left( \frac{1}{1^2} - \frac{1}{n^2} \right)$$

**Balmer series** was discovered in 1885 in the visible region. According to Bohr, in this series, electron jumps to the second orbit ( $m = 2$ ) from higher orbits ( $n = 3,$

4, 5...). The frequencies of various spectral lines of the series are given by

$$\nu_{2n} = \frac{R}{h} \left( \frac{1}{2^2} - \frac{1}{n^2} \right); n > 2$$

**Paschen series** was discovered in 1908 in the near infra-red region. The existence of this series can be explained by assuming that electrons jump to third orbit ( $m = 3$ ) from higher orbits ( $n = 4, 5, 6...$ ). The frequencies of various spectral lines in the region are given by the following equation:

$$\nu_{3n} = \frac{R}{h} \left( \frac{1}{3^2} - \frac{1}{n^2} \right); n > 3$$

**Brackett series** was discovered in mid infra-red region. In this series, electrons jump to fourth orbit ( $m = 4$ ) from higher orbits ( $n = 5, 6...$ ). Therefore, the frequencies of various spectral lines in the region are given by

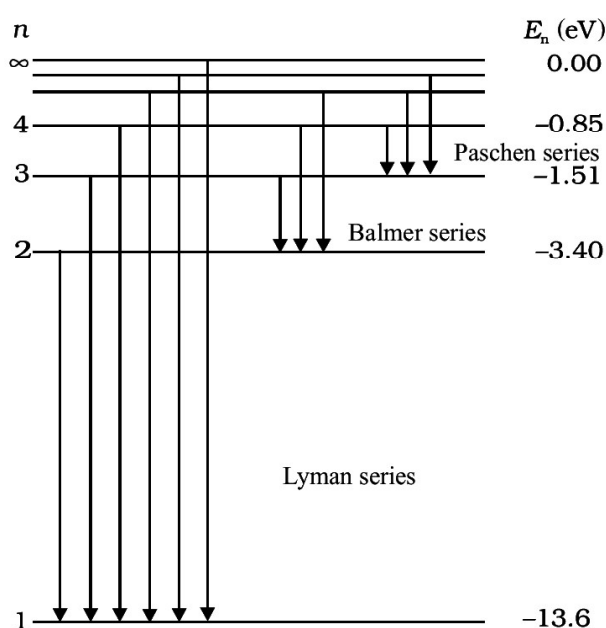
$$\nu_{4n} = \frac{R}{h} \left( \frac{1}{4^2} - \frac{1}{n^2} \right); n > 4$$

**Pfund series** was discovered in far infra-red region. According to Bohr, this series is obtained when electron jumps to fifth orbit ( $m = 5$ ) from any higher orbit ( $n = 6, 7, ...$ ).

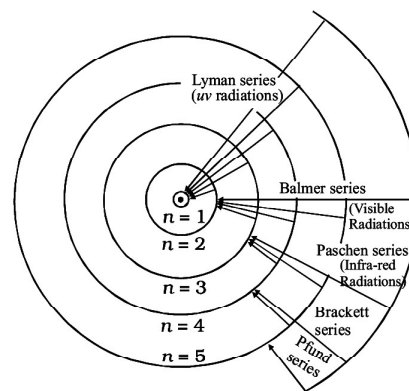
The frequencies of various spectral lines of the series are explained by the following equation

$$\nu_{5n} = \frac{R}{h} \left( \frac{1}{5^2} - \frac{1}{n^2} \right); n > 5$$

The ingenuity of Bohr's model lies in the fact that it not only explained the already known spectrum but also predicted the existence of a number of series. Transition of the electrons from higher orbits to lower orbits showing emission of different series of spectral lines is shown in Fig. 26.8.



**Fig. 26.7 :** Energy Level diagram showing emission of various spectral series in hydrogen



**Fig. 26.8 :** Permitted orbits in an atom of hydrogen and transitions leading to spectral lines of various series



**Intext Questions 26.3**

1. The negative total energy of an orbital electron means that \_\_\_\_\_
2. When the electron in hydrogen atom jumps from 4<sup>th</sup> orbit to the ground state, the number of spectral lines emitted will be \_\_\_\_\_
3. Lyman series of spectral lines are emitted when electron jump from higher orbits to the \_\_\_\_\_ orbit.
4. Which physical property of electron was quantized by Bohr?
5. An electron jumps from third orbit to first orbit. The change in angular momentum of an electron is \_\_\_\_\_

**WHAT YOU HAVE LEARNT**

- Rutherford's scattering experiment indicated the presence of small central region inside the atom and all the positive charge of the mass of the atom is concentrated. The region was named as the nucleus.
- Electrons revolve around the nucleus and total negative charge is equal to the total positive charge of the nucleus.
- Rutherford's model of atom could not explain completely about the stability of the atom and the electromagnetic radiation emitted by the atoms.
- A satisfactory model of an atom was suggested by Niels Bohr based on four postulates.
- Permissible orbits for electrons are those for which angular momentum  $(I\omega) = nh/2\pi$
- Emission (absorption) of energy takes place when electron jumps from a higher orbit to a lower orbit (from a lower to a higher orbit).
- The radii of the permitted orbits in which the electron is free to revolve around the nucleus of the hydrogen atom are given by  $a_n = \frac{n^2 h^2 \epsilon_0}{Z e^2 m \pi}$

For hydrogen atom, the radius of the first permitted orbit is  $a = 0.53 \text{ \AA}$ .

- The energy of the electron in the n<sup>th</sup> orbit of the hydrogen atom is given by

$$E_n = -\frac{e^4 m Z^2}{8 h^2 \epsilon_0^2 n^2}$$

The negative sign of total energy indicates that the electron is bound to the nucleus.

- The frequency of the photon emitted when the electron moves from the energy level  $E_i$  to  $E_f$  is given

$$\nu_{mn} = \frac{R}{h} \left[ \frac{1}{m^2} - \frac{1}{n^2} \right]$$

### TERMINAL EXERCISE

1. In Rutherford's scattering experiment why do most of the  $\alpha$ -particles pass straight through the target foil?
2. In Rutherford's  $\alpha$ -particle scattering experiment, what observation led him to predict the existence of nucleus?
3. Why did Rutherford assume that electrons revolve in circular orbits around the nucleus?
4. What is the ratio of the energies of the hydrogen atom in its first excited state to that its second excited state?
5. What is the SI unit of Rydberg's constant?
6. The Rydberg constant for hydrogen is  $1096700 \text{ m}^{-1}$ . Calculate the short and long wavelength limits of Lyman series.
7. How many times does the electron of H-atom go round the first orbit in 1s?
8. Describe Rutherford's scattering experiment and discuss its findings and limitations.
9. State the postulates of Bohr's model of atom.
10. Derive an expression of the energy of the electron in the  $n^{\text{th}}$  orbit of hydrogen atom.
11. Calculate the radius of the third and fourth permitted orbits of electron in the hydrogen atom.
12. The energy transition in H-atom occurs from  $n = 3$  to  $n = 2$  energy level. Given  $R = 1.097 \times 10^7 \text{ m}^{-1}$ .
  - (i) What is the wavelength of the emitted radiations?
  - (ii) Will this radiation lie in the range of visible light?
  - (iii) To which spectral series does this transition belong

### ANSWERS TO INTEXT QUESTIONS

#### 26.1

1. Alpha rays
2. Protons
3. It could not explain the large angle scattering of particles from the gold foils as observed by Rutherford.

#### 26.2

1. Bohr's first postulate is from classical physics and the remaining three are from quantum physics.
2. Because the orbits are stationary.

## TOSS

3. (i) Electron falls from higher to lower energy state.  
(ii) Electron is excited to some higher energy state.
4.  $E_1 = -13.6\text{eV}$ ,  $E_2 = -3.4\text{ eV}$ ,  $E_3 = -1.51\text{ eV}$ .
5.  $\lambda = hc / (E_1 - E_0)$
6.  $n^2$
7.  $e^4$

## 26.3

1. It is bounded in the nucleus
2. Number of spectral lines emitted =  $\frac{1}{2} n (n - 1) = \frac{1}{2} (4) (4 - 1) = 6$
3. First
4. Angular momentum of revolving electron.
5.  $h / \pi$

## ANSWERS TO TERMINAL EXERCISE

4. 9 : 4
6.  $\lambda_s = 911.4\text{ \AA}$ ,  $\lambda_l = 1215\text{ \AA}$
7.  $6.57 \times 10^{15}\text{ Hz}$ .
12. (i)  $6563\text{ \AA}$   
(ii) visible  
(iii) Balmer Series



# NUCLEAR PHYSICS

## INTRODUCTION

We know that atom is the smallest entity that acts as the building block of all matter. It consists of an extremely small *central core*, called the *nucleus*, around which electrons revolve in certain specified orbits. Though nucleus is very tiny, it is *amazingly complex and you may like to know more about it*. Our understanding the physics of nuclei began towards the end of nineteenth century with the chance discovery of the natural phenomenon of radioactivity (disintegration of atomic nuclei to attain stability). This discovery of the following questions has provided us tools to probe the structure of nucleus: (i) what is its size and mass? (ii) What does a nucleus contain? (iii) What forces make its constituent particles club together and why? These investigations opened up very fertile and new avenues of research. A lot of good new physics of the atom began to emerge out and changed the course of developments in a short span of time. You will learn about these now in this chapter.

## OBJECTIVES

After studying this lesson, you should be able to understand

- *the sizes of atomic nuclei;*
- *the nature of forces between nucleons;*
- *terms 'mass defect' and 'binding energy';*
- *binding energy per nucleon curve and the stability of atomic nuclei;*
- *the phenomenon of radioactivity;*
- *the three types of radioactive radiations;*
- *the growth and decay of radioactivity in a sample;*
- *half-life, and decay constant of a radioactive substance;*
- *uses of radioactivity in various fields;*
- *conservation laws for nuclear reactions;*
- *nuclear chain reaction, controlled and uncontrolled fission chain reactions;*
- *the working of a nuclear reactor;*
- *the mechanism of production of energy in stars.*

## 27.1) THE ATOMIC NUCLEUS

After the discovery of nucleus in an atom by Rutherford in 1911, physicists tried to study as to what resides inside the nucleus. The discovery of neutron by James Chadwick in 1932 gave an impetus to these searches as it clearly suggested to the scientific world that the building blocks of the nucleus are the protons and the neutrons.

### 27.1.1 Charge and mass of the nucleus

The atomic nucleus contains two types of particles, *protons and neutrons*. While *protons* are *positively charged*, *neutrons* are *neutral*. The *electrons*, which revolve in certain specified orbits around the nucleus, are *negatively charged* particles. The magnitude of charge on a proton in a nucleus is exactly equal to the magnitude of charge on an electron. Further, the number of protons in a nucleus is also equal to the number of electrons so that the atom as a whole is electrically neutral. Neutrons and protons are collectively referred to as *nucleons*. Their combined number in a nucleus, that is the number of nucleons, is called the *mass number*. It is denoted by  $A$ . The number of protons in a nucleus (or the number of electrons in an atom) is called the atomic number. It is denoted by  $Z$ . The number of neutrons in a nucleus is usually denoted by  $N = A - Z$ . Usually  $N > Z$ . The difference  $(N - Z)$  increases as  $A$  increases. Note that for a lithium nucleus containing 3 protons and 4 neutrons, the atomic number  $Z$  is 3, and the mass number  $A$  is 7.

Protons are slightly lighter than neutrons and almost the entire mass of an atom is concentrated in its nucleus. The mass of a nucleus is nearly equal to the product of mass number ( $A$ ) and the mass of a proton (or that of a neutron). Since mass of a proton is  $1.67 \times 10^{-27}$  kg, and mass number ( $A$ ) lies between 1 and 240 for most nuclei, the masses of nuclei vary roughly between  $1.67 \times 10^{-27}$  kg and  $4.0 \times 10^{-25}$  kg.

The charge of a nucleus is equal to  $Ze$ , where  $e$  is the fundamental unit of charge (that is the magnitude of charge on an electron,  $1.6 \times 10^{-19}$ C). For naturally occurring nuclei,  $Z$  varies from 1 to 92, while for transuranic elements (i.e. the artificially produced elements),  $Z$  varies from 93 to 105.

### 27.1.2 Size of the nucleus

The sizes of atomic nuclei are usually quoted in terms of their radii. Many nuclei are nearly spherical in shape and the radius  $R$  is given approximately by the formula:  $R = r_0 A^{1/3}$ . Here  $r_0$  is the unit nuclear radius and its numerical value is equal to 1.2 Fermi. The term Fermi is a unit of a radius, given in honour of famous physicist Enrico Fermi. It is equal to  $10^{-15}$  m. The radius of the lightest nucleus (hydrogen) is thus about 1.2 fm, as  $A$  for hydrogen is one. The radius of the heaviest naturally occurring nucleus (uranium) is approximately 7.5 fm, as  $A = 238$ . You may note here that since the volume of any spherical object of radius  $r$  is equal to  $(4/3) \pi r^3$ , the volume of a nucleus is proportional to  $A$ , the mass number. Another important point is that the nucleus and of the atom is approximately  $10^{-15}$  m and  $10^{-10}$  m, respectively. The volume of an atom is roughly  $10^{+5}$  times the volume of a nucleus.

For example, the mass of hydrogen is  $1.673 \times 10^{-27}$  kg and the radius is  $1.2 \times 10^{-15}$  m. The shape is spherical and the density can be calculated using the following relation:

$$d_H = \frac{M_H}{\frac{4\pi}{3} R_H^3} = \frac{1.673 \times 10^{-27} \text{ kg}}{\frac{4\pi}{3} \times (1.2 \times 10^{-15} \text{ m})^3} = 2.3 \times 10^{17} \text{ kg m}^{-3}.$$

For oxygen,  $R_O = 3 \times 10^{-15} \text{ m}$  and  $M_O = 2.7 \times 10^{-26} \text{ kg}$ , so that

$$d_O = 2.39 \times 10^{17} \text{ kg m}^{-3}$$

That is, the densities of hydrogen and oxygen are of the same order. You may recall that the density of water is  $10^3 \text{ kg m}^{-3}$  and density of mercury is  $13.6 \times 10^3 \text{ kg m}^{-3}$ . It means that nuclear matter is extremely densely packed. To give you an idea of these magnitudes, if our earth were such a densely packed mass ( $= 6 \times 10^{24} \text{ kg}$ ), it would be a sphere of radius 184 m only. Similarly, the radius of nuclear sphere, whose mass will be equal to the mass of our sun will be 10 km!

### 27.1.3 Notation used in nuclear physics

The nucleus of an atom is represented by the chemical symbol of the element, with the  $A$  value as its superscript and  $Z$  value as its subscript; both on the left hand side of the chemical symbol. Thus, if the chemical symbol of an element is, say,  $X$ , its nucleus is represented by  ${}^A_Z X$ . For example, for the nucleus of chlorine, which has 17 protons and 18 neutrons, we write  ${}^{35}_{17} \text{Cl}$ . Note that 35 here is mass number.

The atoms of different elements can have the same mass number, though they may have different number of protons. *Atoms having the same  $A$  value but different  $Z$  values are called Isobars.* Thus, argon with  $A = 40$  and  $Z = 18$  is an isobar of calcium which has  $A = 40$  and  $Z = 20$ . Note that isobars have different chemical properties since these are determined by  $Z$ . *Atoms of the same element having the same  $Z$  value but different  $A$  values are called isotopes.* Thus, chlorine with  $Z = 17$  and  $A = 35$ , and chlorine with  $Z = 17$  and  $A = 37$ , are isotopes of same element, chlorine. Since isotopes have same  $Z$  value, they show identical chemical properties. Note that isotopes differ in the number of neutrons in their nuclei. Atoms having the same number of neutrons in their nuclei are called the isotones. Thus, sodium with  $A = 23$  and  $Z = 11$  is an isotone of magnesium with  $A = 24$  and  $Z = 12$ .

#### Example 27.1

Calculate the number of electrons, protons, nucleons in an atom of  ${}^{238}_{92} \text{U}$ .

**Solution :**

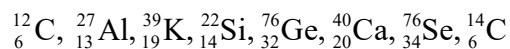
${}^{238}_{92} \text{U}$  symbolizes uranium, which has 92 protons and 238 nucleons. Hence Atomic number  $Z = 92 =$  number of protons

Mass number  $A = 238 =$  number of (protons + neutrons) = Number of nucleons  
Number of neutrons =  $A - Z$

$$\begin{aligned} &= 238 - 92 \\ &= 146. \end{aligned}$$

**Example 27.2**

Select the pairs of Isotopes, Isobars and Isotones in the following list.



**Solution :**

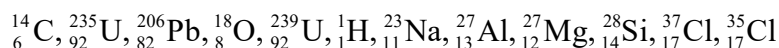
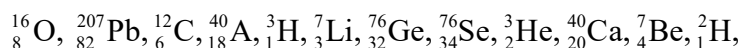
Isotopes – (Same Z - value) :  ${}^{12}_6\text{C}$  and  ${}^{14}_6\text{C}$

Isotones – [Same A - Z values] :  $[{}^{27}_{13}\text{Al}$  and  ${}^{28}_{14}\text{Si}]$ ,  $[{}^{39}_{19}\text{K}$  and  ${}^{40}_{20}\text{Ca}]$

Isobars – (Same A values) :  ${}^{76}_{32}\text{Ge}$  and  ${}^{76}_{34}\text{Se}$

**Intext Questions 27.1**

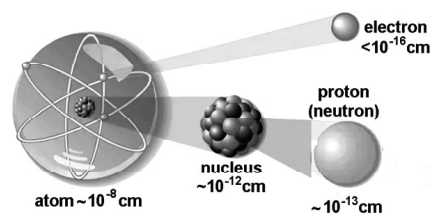
1. Make groups of Isotopes, Isobars and Isotones from the following collection of different atoms:



2. Neutron is ..... than proton.
3. The total number of protons and neutrons in an atom is called the ..... number of that atom.
4. The protons and neutrons together are called .....
5. The number of neutrons in  ${}^{27}_{13}\text{Al}$  = ..... and in  ${}^{28}_{14}\text{Si}$  .....
6. Two atoms are said to belong to different elements if their ..... numbers are different
7. Which number cannot be different in two atoms of the same element .....  
(i) mass number      (ii) atomic number      (iii) neutron number

**27.1.4 Unified Atomic Mass**

It has been experimentally determined that mass of proton ( $m_p$ ) is equal to 1836 times the mass of electron ( $m_e$ ). The mass of neutron ( $m_n$ ) is 1840 times of  $m_e$ . Fig. 27.1 shows size of atom and nucleus. Since the mass of an electron is small compared to the mass of a nucleon, the mass of an atom is effectively due to the mass of its nucleons only. However, the neutron is slightly heavier than the proton, it is, therefore, desirable to choose a standard to express the masses of all the atoms (protons and neutrons). Presently atomic masses are expressed in terms of the actual mass of isotope of carbon. The unit



**Fig. 27.1 :** Size of atom, nucleus and electron

of atomic mass, abbreviated as u, is defined as  $(1/12)^{\text{th}}$  of the actual mass of  $^{12}_6\text{C}$ . We know that the value of the mass of a carbon atom is  $1.99267 \times 10^{-26}$  kg.

$$\begin{aligned} 1u &= (1/12) \times \text{mass of one carbon atom with } A = 12 \\ &= (1/12) \times (1.99267 \times 10^{-26} \text{ kg}) \\ &= 1.660565 \times 10^{-27} \text{ kg} \\ &= 1.66 \times 10^{-27} \text{ kg} \end{aligned}$$

Mass of proton and mass of neutron in terms of u is:

$$m_p = \frac{1.6723 \times 10^{-27}}{1.6606 \times 10^{-27}} u = 1.00727 u$$

$$m_n = \frac{1.6747 \times 10^{-27}}{1.6606 \times 10^{-27}} u = 1.00865 u$$

According to mass-energy relation (Einstein equation), u in terms of Mev is:

$$\begin{aligned} 1u &= (1.66 \times 10^{-27} \text{ kg}) (2.9989 \times 10^8 \text{ ms}^{-1})^2 = 14.92 \times 10^{-11} \text{ J} \\ &= \frac{14.92 \times 10^{-11}}{1.60 \times 10^{-13}} \text{ MeV} = 931.3 \text{ MeV} \end{aligned}$$

### 27.1.5 Mass Defect and Binding Energy

The mass of the nucleus of an atom of any element is always found to be less than the sum of the masses of its constituent nucleons. This difference in mass is called **mass-defect**. For example, the nucleus of deuterium isotope of hydrogen has one proton and one neutron. The measured masses of these particles are  $1.6723 \times 10^{-27}$  kg and  $1.6747 \times 10^{-27}$  kg, respectively. It means that total mass of a proton and a neutron is  $3.34709 \times 10^{-27}$  kg. But the mass of deuterium nucleus is  $3.34313 \times 10^{-27}$  kg. It means that the measured mass of deuterium nucleus is  $3.96242 \times 10^{-30}$  kg less than the measured masses of a proton plus a neutron. So we say that mass defect in the case of deuterium is  $3.96242 \times 10^{-30}$  kg. Let us Z denote it by  $\Delta m$ . Mathematically, for an atom, denoted by , we can write sum of the masses of the nucleons =  $Zm_p + (A-Z)m_n$

$$\therefore \Delta m = [Z.m_p + (A-Z)m_n] - M \quad (27.1)$$

Here, M is the actual mass of nucleus.

Energy equivalent of mass defect is obtained by using mass-energy equivalence relation:

$$BE = \Delta m c^2 \text{ joules} \quad (27.2)$$

For Deuterium

$$\begin{aligned} BE &= (3.96242 \times 10^{-30} \text{ kg}) \times (2.998 \times 10^8 \text{ ms}^{-1})^2 \\ &= 35.164 \times 10^{-14} \text{ kgm}^2\text{s}^{-2} \end{aligned}$$



$$= 3.5164 \times 10^{-13} \text{ J}$$

$$= 2.223 \times 10^6 \text{ eV}$$

$$\text{Since } 1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$$

In other words, *mass defect appears as energy which binds the nucleons together*. This is essentially used up in doing work against the forces which make the nucleons to cling.

Binding Energy per nucleon,  $B = \Delta m c^2 / A$

$$B = \frac{[Zm_p + (A - Z)m_n - M]c^2}{A} \quad (27.3)$$

For carbon,  $Z = 6$  and  $A = 12$ , the binding energy (B) is:

$$B = \frac{[6m_p + 6m_n - 12] \times 931.3}{12} \text{ MeV}$$

$$= 7.41 \text{ MeV}$$

(Here,  $m_p = 1.00727 \text{ u}$  and  $m_n = 1.00865 \text{ u}$ )

The value of B is found to increase to about 8.8 MeV as we move from helium ( $A = 4$ ) to iron ( $A = 56$ ); thereafter it decreases gradually and drops to about 7.6 MeV for uranium ( $A = 238$ ). Fig. 27.2 shows the variation of binding energy per nucleon with mass number.

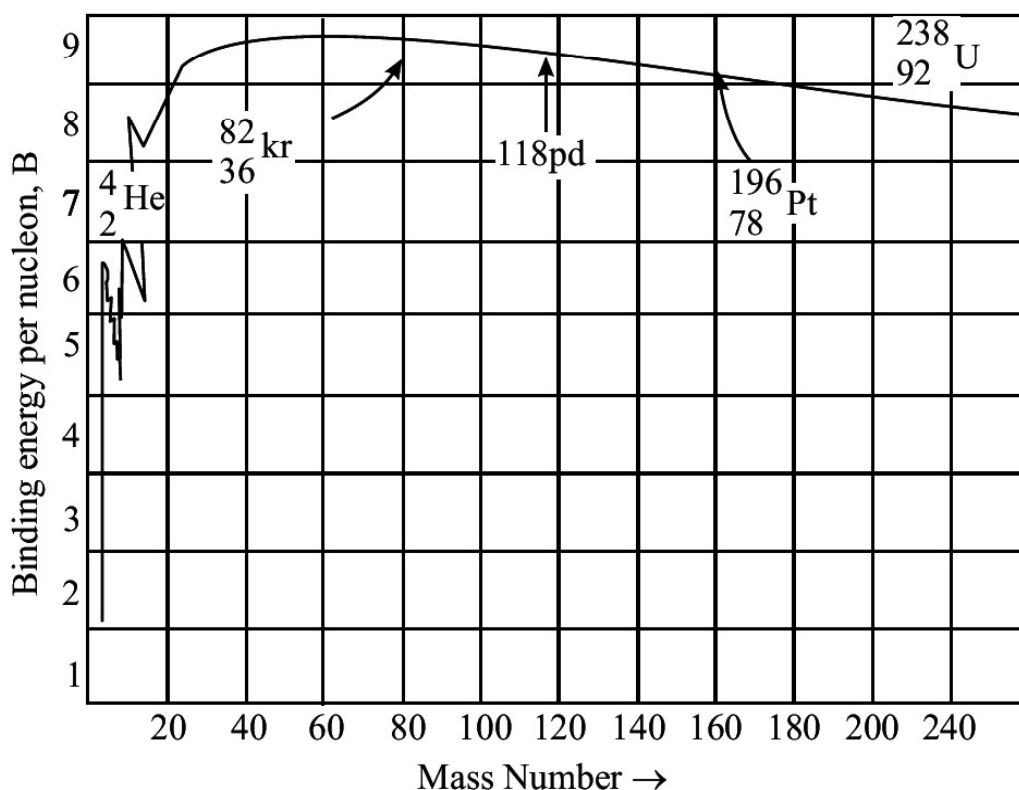


Fig. 27.2 : The variation of binding energy per nucleon with mass number

Note that, binding energy curve shows sharp peaks for  ${}^4_2\text{He}$ ,  ${}^{12}_6\text{C}$ ,  ${}^{16}_8\text{O}$  and  ${}^{20}_9\text{Ne}$ . The binding energy is found to be small for light nuclei with  $A < 20$ . For example, the value of  $B$  for heavy hydrogen ( ${}^2_1\text{H}$ ) is only 1.1 MeV per nucleon. The subsidiary peaks occurring at  ${}^4_2\text{He}$ ,  ${}^{12}_6\text{C}$ ,  ${}^{16}_8\text{O}$  (even-even nuclei i.e. nuclei having even number of protons and even number of neutrons) indicate that these nuclei are more stable than their immediate neighbors. The binding energy per nucleon curve is very useful in explaining the phenomena of nuclear fission and nucleon fusion. This would be discussed at the end of this chapter.

### Example 27.3

Mass of a Boron ( ${}^{10}_5\text{B}$ ) atom is 10.811 u. Calculate its mass in kg.

**Solution :**

Since  $u = 1.660565 \times 10^{-27}$  kg,

$$\begin{aligned} 10.811 \text{ u} &= 10.811 \times 1.660565 \times 10^{-27} \text{ kg} \\ &= 17.952368 \times 10^{-27} \text{ kg} \end{aligned}$$

### Intext Questions 27.2

- The mass of the nucleus of  ${}^7_3\text{Li}$  atom is 6.01513 u. Calculate mass defect and binding energy per nucleon. Take,  $m_p = 1.00727$  u.;  $m_n = 1.00865$  u and  $1 \text{ u} = 931 \text{ MeV}$ .
- Calculate the radius of the nucleus of  ${}^8_4\text{Be}$  atom.

[Use  $R = r_0 A^{1/3}$ ;  $r_0 = 1.2 \times 10^{-15} \text{ m}$ ]

## 27.2) NUCLEAR FORCE

Once physicists ask themselves: How do nucleons cling together? In other words, what is the nature of force that binds nucleons? Before answering this question, let us think, how gravitation and electromagnetic interactions explain most of the observed classical facts. The nuclear force explains the extremely small size of the nucleus, where protons and neutrons are closely packed. The fact suggests that forces should be strong, short range and attractive. These attractive forces cannot have electrostatic origin because electrostatic forces between protons are repulsive. If electrostatic force is operative, then the nucleons would fly away, which is contrary to the existence. Moreover, the forces between nucleons are responsible for the large binding energy per nucleon (nearly 8 MeV). Here, it should consider that the gravitational force is very weak to account for the powerful attractive forces between nucleons. If the magnitude of nucleon- nucleon force is taken to be unity, the gravitational force would be of the order of  $10^{-39}$ . We may, therefore, conclude that the purely attractive forces between nucleons are of a new type with no analogy whatsoever with the forces known in the realm of classical physics. This new attractive force is called **nuclear force**.

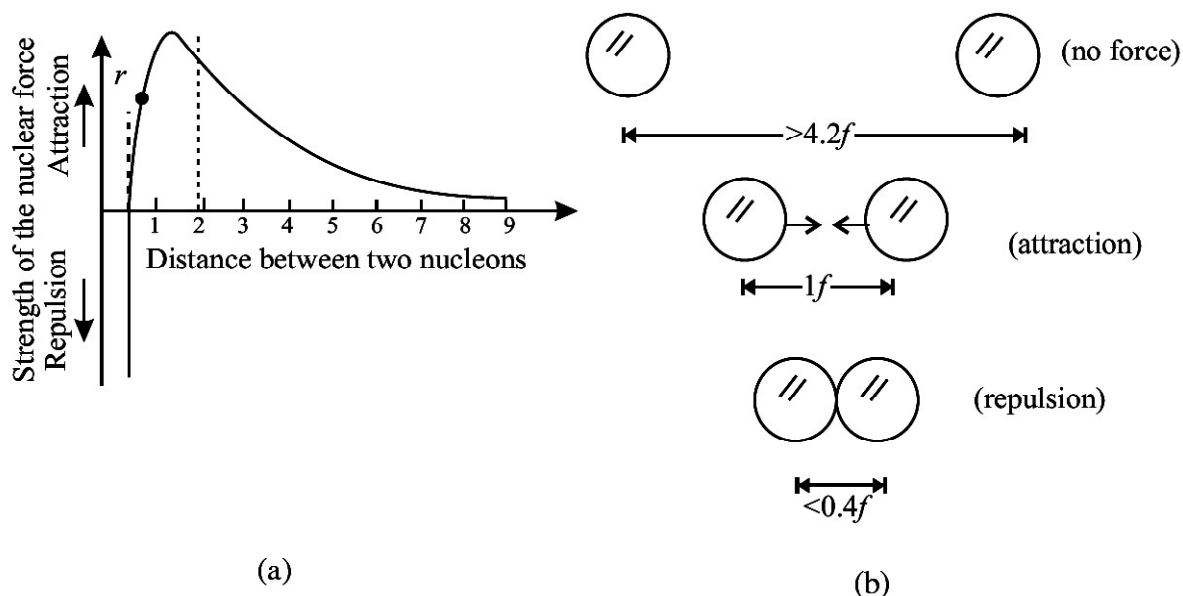
### 27.2.1 Characteristic Properties

You may recall that the gravitational as well as electrostatic forces obey inverse square law. However, the nucleons are very densely packed and the nuclear force that holds the nucleons together in a nucleus must exist between the neighboring nucleons, operating over a short distances ( $\sim 10^{-15}\text{m}$ ).

These nuclear forces must account for the attractive force between:

- a proton and a neutron;
- two protons; and
- two neutrons.

Since binding energy per nucleon,  $B$  is the same, irrespective of the mix of neutrons and protons in the nucleus. That is, nuclear force is *charge independent*. Moreover, the nuclear force shows the property of *saturation*, which means that nucleons show only limited attraction. If nuclear forces had only attractive character, nucleons should have coalesced under their influence. But we all know that the average separation between nucleons is constant, resulting in a nuclear volume proportional to the total number of nucleons. The possible explanation is that nuclear forces exhibit attractive character only so long as nucleons are separated through a certain critical distance. For distances less than this critical value, the character of nuclear forces changes abruptly; attraction should change to repulsion. (You should not confuse this repulsion with electrostatic repulsion.) These qualitative aspects of nuclear forces are shown in Fig. 27.3.



**Fig. 27.3 :** (a) Typical variation of nuclear forces with distance, and (b) effect of inter-nuclear distance on the force between nucleons.

## 27.3) RADIOACTIVITY AND ITS DISCOVERY

What is the age of our earth? How do geologists estimate the age of rocks and fossils found during excavations? What is radio-therapy which is used to treat malignant cells? The answers to all these interesting and useful questions are inherent in the study of radioactivity; a natural phenomenon in which atoms emit radiations to attain stability. Though it was discovered by chance, it has opened flood gates for new physics. Nuclear physics finds wide use in industry, agriculture and medical care. Let us learn about it now.

### 27.3.1 Discovery

The story of discovery of radioactivity is very interesting. In 1896, French physicist A.H. Becquerel was working on the phenomenon of fluorescence (in which some substances emit visible light when they are exposed to ultra-violet radiations). In one of the drawers of his desk, he had kept a collection of various minerals, besides several unopened boxes of photographic plates. Somehow, the collection of minerals remained untouched for a considerable period of time. One day Becquerel used one of the boxes of photographic plates to photograph something. When he developed the plates, he was disappointed to find that they were badly fogged as if previously exposed to light. He tried the other boxes of photographic plates and found them also in the same poor condition. He could not understand as to why plates were fogged because all the boxes were sealed and the plates inside were wrapped with thick black paper.

Becquerel was puzzled and investigated the situation further. He found that uranium placed in his desk had done the damage and concluded that there must be some new type of radiation originating from the uranium salt. This radiation was named *Becquerel rays* and the phenomenon of emission of this radiation was named as *radioactivity*. The elements exhibiting this phenomenon were called *radioactive elements*. Soon after this discovery, and based on an exhaustive study, Madame Marie Curie along with her husband Pierre Curie, isolated an element from uranium ore by a painstaking method known as chemical fractionating. This new element, which was a million times richer in the mysterious rays than uranium, was given the name radium. Another radioactive element discovered by Madam Curie was named polonium in honour of her native country-Poland.

### 27.3.2 Nature of Radiations

In 1899, Lord Rutherford, a British physicist, analyzed the Becquerel rays emitted by radioactive elements. He established the existence of two distinct components:  $\alpha$ -particles and  $\beta$ -rays. The existence of third radiation – gamma rays – was established by P. Villars

We know that nuclei of all atoms contain positively charged protons, which repel each other strongly due to electrostatic repulsion. To overcome this repulsion, neutrons in the nuclei act as glue. But in case of heavier nuclei, this electrostatic repulsion is so strong that even the addition of neutrons is not able to keep the nuclei stable. To achieve stability, such nuclei disintegrate spontaneously by emitting  $\alpha$  and  $\beta$  particles along with  $\gamma$ -rays as shown in Fig. 27.4. So, we can say that in natural radioactivity,  $\alpha$ ,  $\beta$  and  $\gamma$ -rays are emitted. The emitted radiation is called the **radioactive radiation** and the process of disintegration (break-up) of atomic nuclei (by emitting  $\alpha$ ,  $\beta$  and  $\gamma$ -rays) is called **radioactive decay**. Sometimes,

the break-up can be induced by bombarding stable nuclei with other light particles (like neutron and protons). It is then called **artificial radio-activity**. The characteristic features of this phenomenon are that it is spontaneous and in the case of  $\alpha$  or  $\beta$  emission, a new nucleus belonging to a new element is formed. That is, one element gets converted into another element. This is thus a nuclear disintegration phenomenon and suggests the possibility of mutation of new nuclei. Let us first study the characteristic properties of  $\alpha$ ,  $\beta$ , and  $\gamma$  radiations.

### (i) $\alpha$ -particles

Alpha particles are helium nuclei ( ${}^2_4\text{He}$ ) and consist of two protons and two neutrons.

Detailed studies of these particles revealed the following properties:

- Being charged particles, they get deflected in electric and magnetic fields.
- They produce fluorescence in substances like zinc sulphide and barium platino cyanide, affect a photographic plate, can induce radioactivity in certain elements and produce nuclear reactions.
- They have great ionizing power. A single particle in its journey through a gas can ionize thousands of gas atoms before being absorbed.
- They have little penetration power through solid substances, and get scattered by thin foils of metals. They can be stopped by 0.02 mm thick aluminum sheet.
- The energies of a particles emitted from a radioactive substance is a characteristic of the emitting nucleus. This corresponds to a variation in their velocity from  $1.4 \times 10^7 \text{ ms}^{-1}$  to  $2.05 \times 10^7 \text{ ms}^{-1}$ .

### (ii) $\beta$ -particles

$\beta$ -Particles can be both positively and negatively charged. They originate in the nucleus in the process of conversion of a neutron into a proton, and vice versa. Further studies of  $\beta$ -particles, have revealed the following properties.

- Being charged particles, they get deflected by electric and magnetic fields.
- They produce fluorescence in materials like zinc-sulphide and barium plationcynide  $[\text{Pt}(\text{CN})_4]^{2-}$ ; and affect photographic plates.
- They can ionize gas atoms but to a much smaller extent than the  $\alpha$ -particles.
- Negatively charged  $\beta$ -particles can pass through few mm of aluminum sheets. They are about 100 times more penetrating than  $\alpha$ -particles.
- Average energies of negative  $\beta$ -particles vary between 2 MeV and 3MeV. Due to their small mass, their velocities vary in range from  $0.33c$  to  $0.988c$ , where  $c$  is velocity of light.

### (iii) $\gamma$ -rays

They are electromagnetic waves of high frequency, and as such highly energetic. They are characterized with the following properties:

- They do not get deflected by electric or magnetic fields.
- They travel with velocity of light in free space.
- Their penetration power is more than that of  $\alpha$  and  $\beta$ -particles;  $\gamma$ -rays can penetrate through several centimeters of iron and lead sheets.
- They have ionizing power that is smaller compared to that of  $\alpha$  and  $\beta$ -particles.
- They can produce fluorescence in materials and affect a photographic plate.
- They knock out electrons from the metal surfaces on which they fall and heat up the surface. Hard  $\gamma$ -rays (i.e. high energy  $\gamma$ -rays) are used in radio therapy of malignant cells.

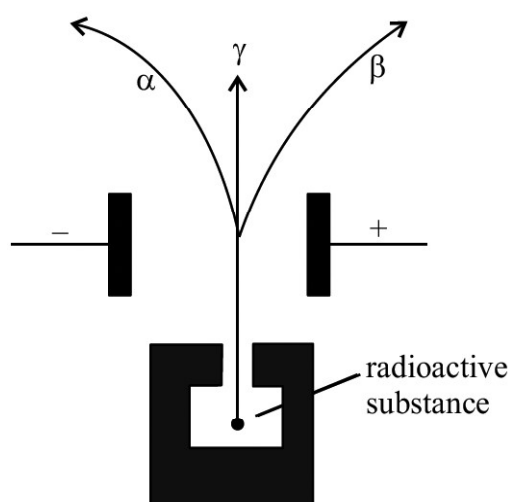
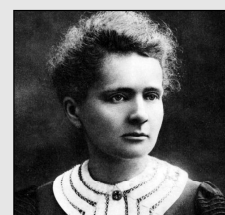


Fig. 27.4 : Emission of  $\alpha$ ,  $\beta$  and  $\gamma$  radiations

### Marie Curie (1867–1934)

Marie Curie shared the 1903 Nobel prize in physics with Henri Becquerel and her husband Pierre Curie for her studies in the field of radioactivity. She was the first person in the world to receive two Nobel prizes; the other Nobel prize she received was in chemistry in 1911. Later her daughter Joliot also won the Nobel Prize in chemistry for her discovery of artificial radioactivity.

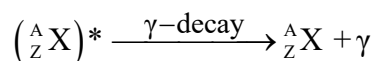
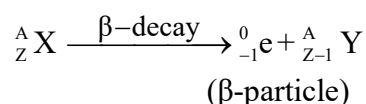
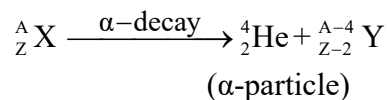


### 27.3.3 Radioactive Decay

In any radioactive decay, spontaneous emission consists of either a single  $\alpha$ -particle or a  $\beta$ -particle. The emission of  $\alpha$ -particle from a radioactive nucleus (called *parent nucleus*) changes it into a new nucleus (new element is called *daughter nucleus*) with its atomic number decreased by two and its mass number decreased by four. Similarly, emission of a  $\beta$  particle changes the parent nucleus into a daughter nucleus with its atomic number

increased by unity (if it is  $\beta^-$  emission) but its mass number remains unchanged. The emission of  $\gamma$ -rays does not change the atomic number or the mass number of the parent nucleus and hence no new nucleus is formed.

Note that in any nuclear disintegration, the charge number (Z) and the mass number (A) are always conserved. Thus for any radioactive nucleus, denoted by X, the nuclear transformations may be written as:



(Here, \* symbol indicates for the excited state)

### 27.3.4 Law of Radioactive Decay

We now know that if we have a given amount of radioisotope, it will gradually decrease with time due to disintegrations. The law describing radioactive decay is very simple. The rate of radioactive disintegration is independent of external factors such as temperature, pressure etc. and depends only on the law of chance. It states that *the number of radioactive atoms disintegrating per second is proportional to the number of radioactive atoms present at that instant of time*. This is called *law of radioactive decay*.

Let  $N_0$  be the number of radioactive atoms, at  $t = 0$ , and  $N(t)$  be the number of radioactive atoms at time  $t$ . If  $dN$  denotes the number of atoms that decay in time  $dt$ , then  $(N - dN)$  signifies the number of radioactive atoms at time  $(t + dt)$ . Hence, rate of decay:

$$\frac{dN(t)}{dt} \propto N,$$

$$\frac{dN(t)}{dt} = -\lambda N(t) \quad (27.4)$$

Where  $\lambda$  denotes decay constant, which is characteristic of the radioactive substance undergoing decay. The negative sign signifies that the number of nuclei decreases with time. This relation can be rearranged as

$$\lambda = -\frac{1}{N(t)} \frac{dN(t)}{dt} \quad (27.5)$$

*Thus, decay constant ( $\lambda$ ) may be defined as the ratio of the instantaneous rate of disintegration to the number of radioactive atoms present at that instant.*

The law of decay is sometimes also expressed in exponential form and is also called the *law of exponential decay*. To obtain the exponential form, we rewrite Eqn. (27.4)

$$\frac{dN(t)}{N(t)} = -\lambda dt$$

On integrating both sides,

$$\ln N(t) = -\lambda t + k$$

$$\text{At } t = 0, N(t) = N_0$$

$$\therefore k = \ln N_0$$

Hence

$$\ln N(t) - \ln N_0 = -\lambda t$$

$$\text{or } \ln \left( \frac{N(t)}{N_0} \right) = -\lambda t$$

$$\text{Finally, } N(t) = N_0 \exp(-\lambda t), \text{ (on taking antilog)} \quad (27.6)$$

The most important conclusion from this law is that  $N$  will become zero only when  $t = \infty$ . Thus, no radioactive element will disappear completely even after a very long time. The radioactive decay law clearly shows that even if the number of atoms  $N_0$  for different radioactive elements is same initially, at a later time they will have different values of  $N(t)$  due to different values of their decay constants ( $\lambda$ ).

### Units of Disintegration

The decay constant is measured in units of per second. The *activity* of a radioactive substance at any instant of time is measured by its rate of disintegration. Its SI unit has been named Becquerel:

$$1 \text{ Becquerel (Bq)} = 1 \text{ disintegration per second.}$$

Another unit of the decay constant is *curie*.

$$1 \text{ curie (Ci)} = 3.7 \times 10^{10} \text{ disintegrations per second.}$$

The rate of disintegration of radium (Ra) measured per second per gram. Yet another unit is 'Rutherford' (Rd) :

$$1 \text{ Rd} = 10^6 \text{ disintegrations per second.}$$

### 27.3.5 Half Life ( $T_{1/2}$ )

*The half-life ( $T_{1/2}$ ) of any radioactive element is defined as the time in which the number of parent radioactive atoms decreases to half of the initial number.*

By definition, at  $t = T_{1/2}$ ,  $N = N_0/2$ . Therefore Eqn. 27.6 can be expressed as,



$$\begin{aligned}
 N_0/2 &= N_0 \exp(-\lambda T_{1/2}) \\
 \text{or } \lambda T_{1/2} &= \log_e 2 \\
 T_{1/2} &= \frac{\log_e 2}{\lambda} \\
 &= \frac{2.303 \times \log_{10} 2}{\lambda} \\
 &= \frac{0.693}{\lambda}
 \end{aligned} \tag{27.7}$$

Thus, half-life of any radioactive substance is inversely proportional to its decay constant and is a characteristic property of the radioactive nucleus. The half-life of  ${}^{14}_6\text{C}$  (radioactive carbon) is 5730 years. This means that one gram of  ${}^{14}_6\text{C}$  will be reduced to 0.5 g in 5730 years. This number will be further reduced to  $0.5 / 2 = 0.25$  g in another 5730 years.

i.e. in a total period of 11460 years. Refer to Fig. 27.5 to see how a radioactive sample decays with time.

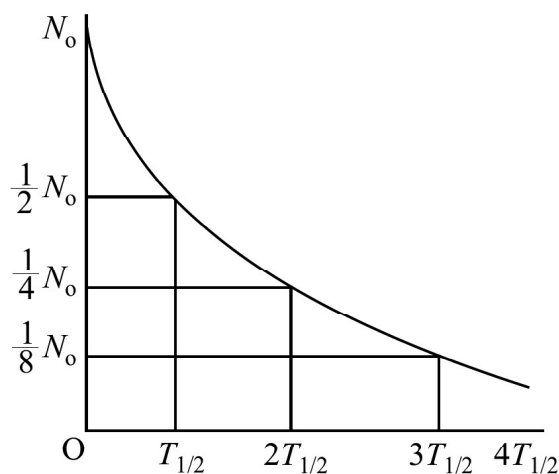


Fig. 27.5 : Radioactive decay curve

#### Example 27.4

An animal fossil obtained in the Mohenjo-Daro excavation shows an activity of 9 decays per minute per gram of carbon. Estimate the age of the Indus Valley civilization. Given the activity of in a living specimen of similar animal is 15 decays per minute per gram, and half-life of  ${}^{14}_6\text{C}$  is 5730 years.

#### Solution :

${}^{14}_6\text{C}$  is radioactive isotope of carbon. It remains in fixed percentage in the living species. However, on death, the percentage of  ${}^{14}_6\text{C}$  starts decreasing due to radioactive decay. Using radioactive decay law, we can write

$$\begin{aligned}
 N(t) &= N_0 \exp(-\lambda t) \\
 \text{so that } N/N_0 &= \exp(-\lambda t) \\
 \text{or } 9/15 &= \exp(-\lambda t) \\
 \text{or } \log_e(9/15) &= -\lambda t
 \end{aligned}$$

or 
$$\log_e \left( \frac{15}{9} \right) = \lambda t$$

which gives 
$$t = 1/\lambda [\log_e (15/9)]$$

Here  $T_{1/12} = 0.693/\lambda = 5730$  years, Therefore,

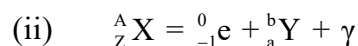
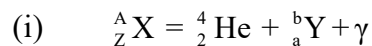
$$t = 2.303 \times (5730/0.693) [\log_{10} 15 - \log_{10} 9]$$

Hence 
$$t = 4224.47 \text{ years.}$$

Thus, the specimen containing carbon-14 existed 4224.47 years ago. Hence the estimated age of Indus valley civilization is 4225 years.

### Intext Questions 27.3

- How can you say that radioactivity is a nuclear disintegration phenomenon?
- Compare the ionizing and penetration powers of  $\alpha$ ,  $\beta$  and  $\gamma$  - radiations.
- Apply the law of conservation of charge and mass numbers to determine the values of a and b in the following decay - equations :



- The half-life of a radioactive substance is 5 years. In how much time, 10 g of this substance will reduce to 2.5 g?

### 27.3.6 Applications of Radioactivity

Radioactivity finds many applications in our everyday life. Some of these are given below.

- In medicine:** In the treatment of cancer (radiotherapy), a radio-active cobalt source which emits  $\alpha$ -rays is used to destroy cancerous cells. The decay of a single radioactive atom can be registered by an instrument placed at a remote location outside a container wall. This high sensitivity is utilized in *tracer technique* as an important tool in medical diagnostics, like the detection of ulcer in any part of the body.
- In agriculture:** By exposing the seeds to controlled  $\gamma$  radiation, we are able to improve the quality and yield of crops, fruits and vegetables. Radiating these seeds before their storage helps in saving from decay.
- In geology:** In estimating the age of old fossils. The normal activity of living carbon containing matter is found to be about 15 decays per minute for every gram of carbon. This activity arises from the small proportion of radioactive carbon-14 present in the atmosphere with the ordinary carbon-12. This isotope ( ${}^{14}_6\text{C}$ ) is taken by plants from the atmosphere and is present in animals that eat plants. Thus, about one part in  $10^8$  radioactive carbons is present in all living beings (all animals and

plants). When the organism is dead, its interaction with the atmosphere (i.e. absorption, which maintains the above equilibrium) ceases and its activity begins to fall. From this, the age of the specimen can be approximately estimated. This is called **carbon-dating** and is the principle of determining the age of old fossils by archeologists. The same technique has been used in estimating the age of earth from the measurements of relative amounts of  $^{238}\text{U}$  and  $^{206}\text{Pb}$  in geological specimens containing uranium ore. Assume that the specimen of ore contained only uranium and no lead at the time of birth of the earth. With the passage of time, uranium decayed into lead. The amount of lead present in any specimen will therefore indicate its age. The present age of the earth, using this method, has been estimated to be about 4 billion years.

4. **In industry:**  $\gamma$ -radiations are used to find the flaws (or imperfections) in the inner structure of heavy machinery. For example, if there is an air bubble inside, the penetration of  $\gamma$ -rays will be more at that point.

## 27.4) NUCLEAR FISSION AND FUSION

It is a known fact that the sun supports life on the earth by continuously providing energy. It has been doing so for the last several billion years and will continue to do so for billions of years to come. What is the source of this huge amount of energy emitted by the sun? This question fascinated human mind always. But now we reliably know that the energy in the core of sun is produced by fusion of hydrogen nuclei into helium at very high temperatures. This is also true of other stars. Imitation of these conditions in a fusion reactor is being highlighted as the ultimate source of all our energy requirements in coming years. Similarly, you must have read about energy security and the role of nuclear energy to produce electricity in our nuclear reactors at Tarapore, Kota, Kaiga, Narora, Kalpakkam and Kakrapar. World witnessed on August 6, 1945, an atom-bomb dropped on Hiroshima, a large city in Japan, destroyed the entire city almost completely in a span of a few seconds and lakhs of lives were lost. It released an energy equivalent to that released by the explosion of a 20,000 ton TNT (tri nitro toluene) bomb and was completely new in human history. Since then, more powerful (atomic, hydrogen and neutron) bombs have been made whose destructive power is equivalent to several Mega tons of TNT. The super powers are said to have stockpiled a large number of such bombs. The destructive power of their stock is so enormous that they can destroy the entire earth several times over. The physical process responsible for such colossal amount of energy is nuclear fission. An aspect of this would be discussed in the following section. Prior to this, the concept of chemical reaction and nuclear reactions are discussed below:

### 27.4.1 Chemical Reaction

All substances are made up of atoms and the electrons in the outermost orbit govern the chemical properties of an element. That is, atoms combine with other atoms or molecules (a group of atoms) and rearrange their valence electrons. This is accompanied by reduction in their potential energy.

**The formation of a new compound molecule due to rearrangement of valence electrons in interacting atoms and molecules with the release or absorption of energy is called a chemical reaction. In this process, the nucleus is not affected at all. Even the electrons in the inner orbits remain unaffected.**

An example of a chemical reaction is the interaction of carbon atoms with oxygen molecules to produce carbon dioxide:



**In this chemical reaction, 4.08 eV energy is released for each reacting carbon atom. It is called the binding energy (B.E) of CO<sub>2</sub> molecule.** Reactions which result in release of energy are said to be *exothermic*. Chemical reactions which require energy to be supplied to be initiated are *endothermic*. For example, if 4.08 eV of energy is given to a CO<sub>2</sub> molecule under suitable conditions, it will break up into its constituents:



As shown in Eqn. (27.8), 4.08 eV energy leaves the system to form CO<sub>2</sub> gas. Therefore, the mass of CO<sub>2</sub> molecule will be less than the total mass of C and O<sub>2</sub> by a mass equivalent of 4.08 eV. The loss of mass  $\Delta m$  can be calculated using the relation  $E = mc^2$  :

$$\Delta m = \frac{4.08 \times 1.602 \times 10^{-19}}{9 \times 10^{16}} = 7.26 \times 10^{-36} \text{ kg} \quad (27.10)$$

Such a small change in mass cannot be detected and we say that the mass is conserved in chemical reactions, though slight change of mass does occur.

The important points to be noted in chemical reactions are:

- Energies of the order of 10 eV are involved.
- Change of mass is of the order of  $10^{-35}$  kg, which is extremely small and we say that the mass is conserved.
- The total number of atoms of each type on the right hand side of the chemical equation is always equal to the total number of atoms of each type on the left hand side.

## 27.4.2 Nuclear Reactions

In nuclear reactions, the nuclei, not electrons, of the reactants interact with each other. They result in the formation of new elements. This process is also called transmutation of nuclei. From the previous lesson, you may recall that in nuclear reactions energies of the order of MeV are involved.

As it was pointed out that the entire positive charge of an atom is concentrated in its nucleus, whose size is of the order of  $10^{-15}$  m. The nucleus is surrounded by electrons revolving in certain specified orbits. These create a strong electrostatic potential barrier (also called the Coulomb barrier) as shown in Fig. 27.6. The Coulomb barrier is about 3 MeV for carbon nuclei and 20 MeV for lead nuclei. It means that a charged projectile aimed at a nucleus will experience strong repulsion by the Coulomb barrier of the target nucleus. If the kinetic energy of projectile is not large enough to penetrate the barrier, it will come

back without producing any nuclear reaction. For a proton to enter a carbon nucleus and produce transmutation, its energy should be more than 3 MeV or so.

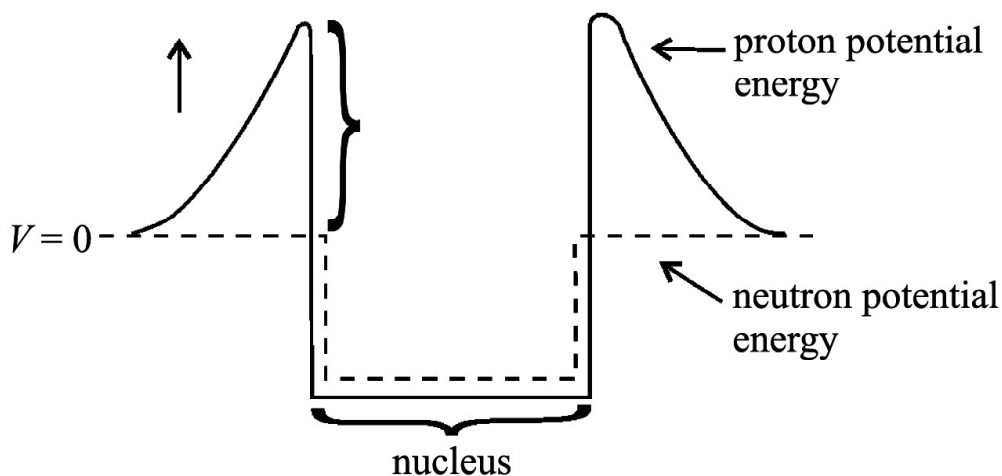


Fig. 27.6 : Proton and neutron potential energies near a nucleus

The phenomenon of nuclear **transmutation** or nuclear reaction was discovered by Lord Rutherford in the year 1919. He bombarded nitrogen gas with high energy  $\alpha$ -particles of energy 7.7 MeV obtained from a polonium source. He observed that nitrogen transformed into oxygen. This change was accompanied by high energy protons:



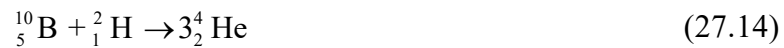
The oxygen nuclei and protons carry away 6.5 MeV. Clearly this reaction can occur if 1.2 MeV energy is supplied from outside. Therefore, it is an endothermic nuclear reaction. When aluminum is bombarded by 7.7 MeV alpha particles from polonium, the following nuclear reaction takes place and 10.7 MeV energy is released:



Here we see that more energy is released than the input energy; it is an exothermic reaction. Note that there is a gain of nearly 3 MeV energy per reaction, which is approximately 700,000 times the energy released in burning of one carbon atom. *But this reaction can't be used for production of energy because out of 125,000 incident alpha particles only one succeeds in producing the reaction.* Hence on the whole, is much more energy spent than produced?

Nuclear reactions can also be produced by protons, deuterons, neutrons and other light nuclei. Of these, *neutrons are the best projectiles for producing nuclear reactions; being neutral particles, they do not experience Coulomb repulsion.* Thus even thermal neutrons (i.e. neutrons having energy 0.0253 eV) can penetrate the target nucleus and produce a nuclear reaction. Some typical examples of nuclear reactions produced by protons, deuterons and neutrons are:





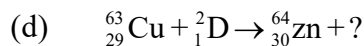
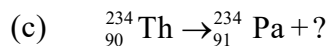
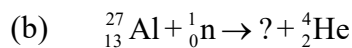
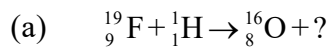
Like chemical reactions, nuclear reactions also follow conservation laws. We state these now.

#### Conservation Laws for Nuclear Reactions

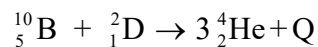
- *The sum of the mass numbers of the reactants is equal to the sum of mass numbers of the products. In Eqn. (27.13), mass number 7 = 3 + 4 = 6 + 1 is conserved.*
- *The sum of atomic numbers of the reactants is equal to the sum of atomic numbers of the products. In Eqn. (27.13), atomic number 4 = 3 + 1 = 2 + 2 is conserved.*
- *Nuclear reactions follow the law of conservation of energy. We know that mass is concentrated form of energy. Therefore the sum of input kinetic energy plus the mass of the reactants is equal to the output kinetic energy plus the mass of the products.*
- *Nuclear reactions follow the law of conservation of momentum, which results in distribution of kinetic energy among various product nuclei*

#### Intext Questions 27.4

1. Complete the following equations of nuclear reaction.



2. Calculate the energy released in the nuclear reaction given below



Given that  $m({}_{5}^{10}\text{B}) = 10.01294\text{u}$ ;  $m({}_{1}^{2}\text{D}) = 2.014103\text{u}$ , and  $m({}_{2}^{4}\text{He}) = 4.002604\text{u}$ .

3.  ${}_{7}^{14}\text{N}$  nucleus, on bombarding with alpha particles, produces  ${}_{8}^{17}\text{O}$ . Write down the reaction equation and calculate the energy released.

Given that  $m({}_{7}^{14}\text{N}) = 14.003014\text{u}$ ;  $m({}_{8}^{17}\text{O}) = 16.999138\text{u}$ ;  $m({}_{2}^{4}\text{He}) = 4.002604\text{u}$ ;  
 $m({}_{1}^{1}\text{H}) = 1.007825\text{u}$  and energy of  $\alpha$  particle = 7.7 MeV.

## 27.5) NUCLEAR FISSION

The story of discovery of fission is very fascinating. In the year 1938, Enrico Fermi, Otto Hahn and others irradiated uranium nuclei with slow neutrons to produce transuranic elements (having  $Z$  greater than 92), which do not occur in nature. When incident neutrons were captured by the uranium nuclei, the neutron-proton ratio increased. In reducing this ratio, it was expected that uranium would become  $\beta$ -active. That is a neutron would essentially behave as if it has changed into a proton resulting in the release of a  $\beta$ -particle and some energy according to the equation:



In this process, a new transuranic element having atomic number 93 was expected to be produced. In fact, Fermi and his co-researchers observed  $\beta$ -activities with half-lives different from any of the known values for heavy elements in the vicinity of uranium. From those observations, they concluded that transuranic elements had been produced. And to identify the element, they carried out chemical analysis but failed.

In the same year, Otto Hahn and Fritz Strausmann carried out a series of experiments and established that barium, an element of intermediate mass number, rather than a transuranic element was one of the products of the reaction and it was accompanied by release of nearly 200 MeV of energy. This result – the product of slow neutron bombardment of uranium was barium – was completely unexpected and defied all knowledge of nuclear physics of that time. These findings were reported in *Nature* in Dec. 1938.

Initially, Lise Meitner and Otto Robert Frisch explained these results on the basis of liquid drop model of nucleus and named this process *nuclear fission* using the analogy with biological cell division. Later on, Bohr and Wheeler calculated the amount of energy released in the process, confirming the physical basis of this model.

### Enrico Fermi (1901 – 1954)

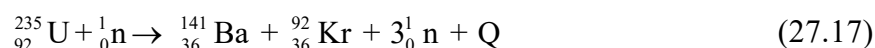
Enrico Fermi, the Italy born physicist, was responsible for peaceful uses of nuclear energy for mankind. He demonstrated that nuclear transformations may occur in any element exposed to stream of neutrons. He achieved self-sustained nuclear fission chain reaction in 1942.

Fermi was only 25 years old when he formulated the Fermi–Dirac statistics, applicable to particles having half integral spin values (called fermions). At the time of his premature death, he was engrossed in theoretical studies of cosmic radiations.



#### 27.5.1 Mechanism of Nuclear Fission

In the year 1939, Bohr and Wheeler developed the theory of fission using the analogy between nuclear forces and the forces which bind molecules in a liquid. They predicted that  ${}_{92}^{235}\text{U}$  was more fissile than  ${}_{92}^{237}\text{U}$ . Refer to Fig. 27.7. It shows the schematics of nuclear fission of by  ${}_{92}^{235}\text{U}$  thermal neutrons according to the equation.



The emitted neutrons have energy of the order of a few MeV, and  $Q$  200MeV. Note that a fission event occurs within  $10^{-17}$ s of neutron capture and fission neutrons are emitted

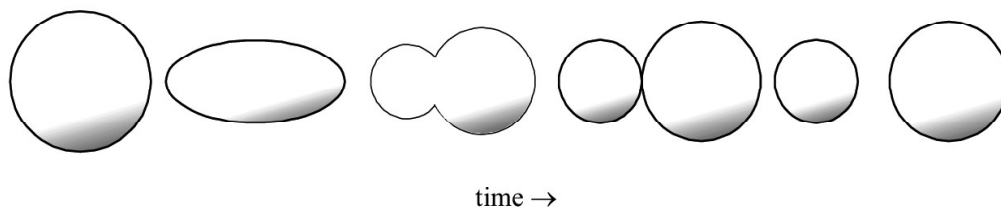


Fig. 27.7 : Nuclear-fission of a nucleus according to the liquid drop model

within about  $10^{-14}$ s of the event. Moreover, the fission fragments are of unequal mass; one being 1.5 to 2 times heavier than the other. Also, Eqn. (27.17) gives only one of the more than 40 different modes in which a  ${}^{235}\text{U}$  nucleus can fission. It means that about 80 different nuclei of intermediate masses are produced in the fission of  ${}^{235}\text{U}$ . The heavier fragments lie in the mass range 125 – 150. The maximum is around 140, whereas the lighter fragments lie in the range 80 – 110 with a maximum around 95. The number of neutrons emitted is either two or three and the average number of neutrons produced per fission of  ${}^{235}\text{U}$  is 2.54

Bohr and Wheeler treated the nucleus as a charged spherically symmetric liquid drop in its equilibrium (lowest energy) state. According to them, when a nucleus captures a thermal neutron, the binding energy (BE) of this neutron, which is 6.8 MeV per atomic mass unit for  ${}^{235}\text{U}$ , is released. This energy excites the nucleus and distorts its shape. While the force of surface tension tries to restore the original shape, the Coulomb force tends to distort it further. As a result, it oscillates between spherical and dumb bell shapes, as shown in Fig. 27.7, depending on the energy of excitation. When the energy gained by the nucleus is large, the amplitude of these oscillations pushes the nucleons into dumbbell shape. When the distance between the two charge centers exceeds a critical value, electrostatic repulsion between them overcomes nuclear surface tension and pushes the nucleus into two parts resulting in fission.

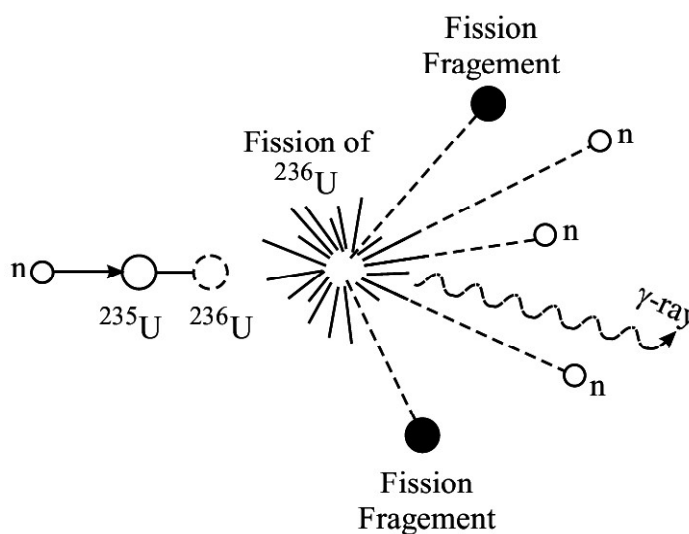


Fig. 27.8 : Nuclear fission

*A substance like  ${}^{235}_{92}\text{U}$  which undergoes fission by thermal neutrons is called a fissile material.* Other fissile materials are  ${}^{233}_{90}\text{Th}$ ,  ${}^{233}_{92}\text{U}$  and  ${}^{239}_{93}\text{Pu}$ . You may note that all these nuclei



have odd mass number and even atomic number. We can estimate the amount of energy released in the fission of  ${}_{92}^{235}\text{U}$  by calculating the mass defect as follows:

**Table - 27.1 Energy Generated in a Nuclear Reaction**

Reactants	mass	Products	Mass
${}^{235}\text{U}$	235.0439 u	${}_{56}^{14}\text{Ba}$	140.9139 u
${}^1_0\text{n}$	1.008665 u	${}_{36}^{92}\text{Kr}$	91.8973 u
		$3 \times \text{Vn}$	3.025995 u
<b>Total mass</b>	236.052565 u	<b>Total mass</b>	235.837195 u
Mass defect	0.21537 u		
Energy released	$0.21537 \times 931 \simeq 200 \text{ MeV}$		

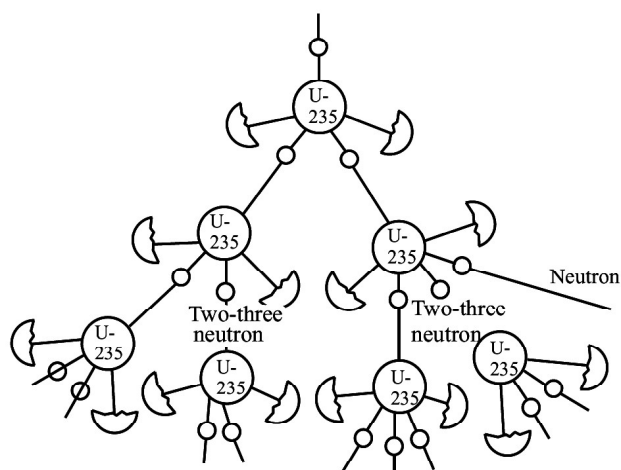
### 27.5.2 Nuclear Chain Reaction

You have now learnt that when a neutron is captured by  ${}_{92}^{235}\text{U}$ , it splits into two fragments and 2-3 neutrons are emitted. These are capable of causing further fissions. This immediately presented the exciting possibility of maintaining a fission chain reaction in which each fission event removes one neutron and replaces that by more than two.

When the rate of production of neutrons equals the rate loss of neutrons, the reaction is said to be *self-sustained*. The device designed to maintain a self-sustained and controlled chain reaction is called a *nuclear reactor*. Nuclear reactors are usually classified according to the purpose for which they are used. So a nuclear power reactor is used to produce electricity and a research reactor is used to produce radioisotopes for medical purposes, carrying out experiments for refinements or applied research. We also categorize nuclear reactors as fast and thermal, depending on the energy of neutrons causing fission. In India, we have thermal power reactors at Tarapore, Narora, Kota, Kaiga, etc. At Kalpakkam, we have developed a fast breeder research reactor. A brief description of nuclear reactor is given below.

### 27.5.3 Nuclear Reactor

Ever since the first nuclear reactor was constructed by Fermi and his co-workers at the university of Chicago USA, a large number of reactors have been built the world over primarily to meet demand for energy. Some countries generate as much as 70% of their



**Fig. 27.9 : Nuclear chain reaction**

total energy from nuclear reactors. In India, the contribution of nuclear energy is only about 2%, but efforts are on to increasing this share. In absolute terms, we are generating about 20,000 MW from nuclear reactors. Nuclear reactors have huge complex structures and great care has to be exercised in designing them. The basic principle of a nuclear power plant is very simple and analogous to any power plant. The heat liberated in fission is used to produce steam at high pressure and high temperature by circulating a coolant, say water, around the fuel. (In a coal fired station, coal is burnt to produce steam. Since one fission event generates about  $7 \times 10^5$  times more energy than that produced in burning one atom of carbon, we can cut down on emission of greenhouse gases substantially by switching over to nuclear energy. However, there are some complex social and political issues with global dimensions that will ultimately decide our ultimate nuclear energy options.)

The steam runs a turbine–generator system to produce electricity. (In research reactors, the heat is discharged into a river or sea. You may have heard about Bhabha Atomic Research Centre at Trombay, Mumbai or Indira Gandhi Atomic Research Centre at Kalpakkam. The heat generated by the research reactors at these centers is discharged into the Arabian sea and the Bay of Bengal, respectively.)

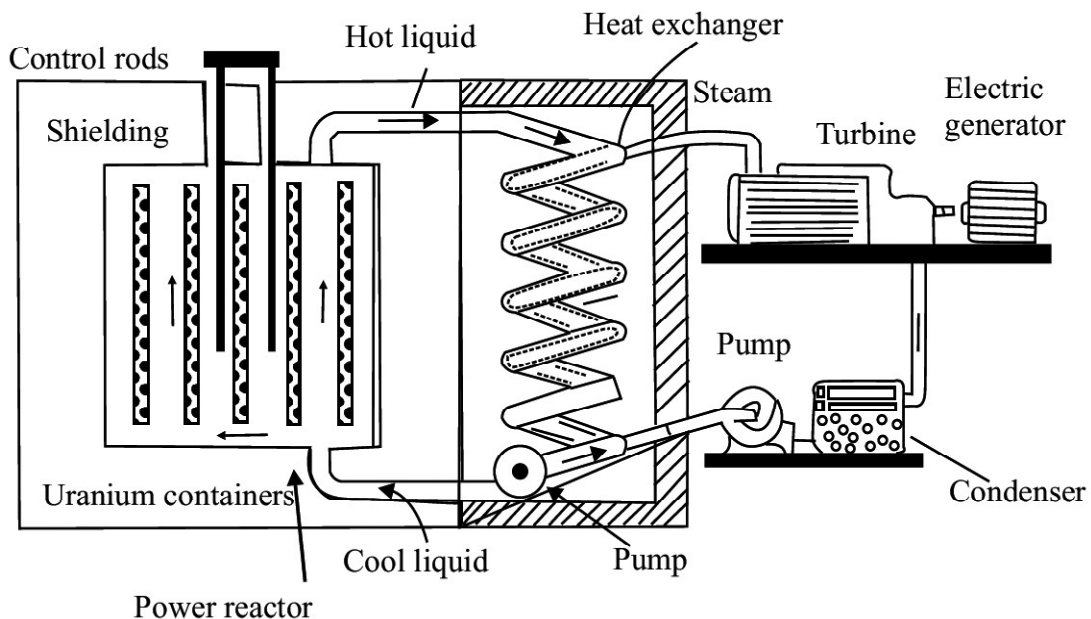


Fig. 27.10 : Schematic diagram of a nuclear reactor

The general features of a reactor are illustrated in Fig. 27.10. All nuclear reactors consist of:

- A *reactor core*, where fission takes place resulting in release of energy. It has fuel rods (embedded in a moderator in a thermal reactors), and *control rods* to maintain the chain reaction at the desired level. *Coolant* is circulated to remove the heat generated in fission. Usually, heavy water or ordinary water is used as coolants and cadmium or boron is used for control rods.
- A *reflector* is put next to the core to stop neutron leakage from the core.

- The whole assembly is placed inside a vessel, called *pressure vessel*. Usually, a few inches thick stainless steel is used for this purpose.
- A thick *shield* is provided to protect the scientists and other personnel working around the reactor from radiations coming from the reactor core. It is usually in the form of a thick concrete wall.
- The entire structure is placed inside a *reactor building*. It is air tight and is maintained at a pressure slightly less than the atmospheric pressure so that no air leaks out of the building.

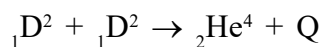
The heat generated inside the reactor core of a reactor due to fission is removed by circulating a *coolant*. The heated coolant is made to give up its heat to a secondary fluid, usually water in a heat exchanger. This generates steam, which is used to drive turbine-generator system to produce electricity in a power plant and discharged into a river/lake/sea in a research reactor.

### Intext Questions 27.5

1. Why does a  ${}_{92}^{238}\text{U}$  nucleus become  $\beta$ -active after absorbing a neutron?
2. Out of  ${}_{92}^{238}\text{U}$ ,  ${}^{141}\text{Ba}$ ,  ${}^{239}\text{Pu}$ , and  ${}_{6}^{12}\text{C}$ , which nucleus is fissile?
3. How much energy is released when  ${}_{92}^{235}\text{U}$  undergoes nuclear fission?

## 27.6) NUCLEAR FUSION

You now know that uranium nucleus can be made to split into lighter nuclei resulting in release of huge amount of energy. You may now ask: Can we combine lighter nuclei to produce energy? To discover answer to this question, refer to the binding energy per nucleon (BE/A) curve (Fig. 27.2). You will note that binding energy per nucleon increases as we go from hydrogen to helium. It means that helium is more stable than hydrogen. Consider the following reaction:



You can easily calculate the B.E of reactants and products using Table-27.2:

**Table - 27.2 : Binding Energy per nucleon (BE/A) of some light nuclei**

Nucleon	BE/A (in MeV)
${}^3\text{D}$	1.11
${}^3\text{T}$	2.827
${}^3\text{He}$	2.573
${}^4\text{He}$	7.074
${}^6\text{Li}$	5.332
${}^7\text{Li}$	6.541

You can easily calculate the B.E of reactants and products:

$$\text{Total B.E of reactants, } BE_1 = 2 \times 2.22 = 4.44 \text{ MeV}$$

$$\text{Total B.E of products, } BE_2 = 28.295 \text{ MeV}$$

$$Q = (BE_2 - BE_1) \simeq 24 \text{ MeV}$$

Therefore, the energy released per nucleon in this reaction is  $24/4 = 6 \text{ MeV}$ , which is nearly seven times the energy released per nucleon ( $200/238 = 0.83 \text{ MeV}$ ) in a nuclear fission event.

***The process in which two light nuclei combine to form a heavier nucleus is called nuclear fusion.***

Fusion process presents itself as a more viable energy option. However, the process of fusion is more difficult to achieve than nuclear fission because both the deuterons are positively charged. When we try to bring them together to fuse into one nucleus, they repel each other very strongly and the reaction is ordinarily impossible.

To achieve this reaction, the deuterons have to be heated to nearly 10 million kelvin so that they acquire sufficient kinetic energy to overcome repulsion before they collide to fuse into helium nucleus. But the problems associated with maintaining such high temperatures continuously and containing the reactants together has not yet been solved fully. The controlled thermonuclear reaction necessary for harnessing this source of energy is however not far now.

Almost inexhaustible amount of deuterium (heavy hydrogen) is present in the ocean. Once we begin to harness this source, our energy problem should be solved forever. We will get endless supply of cheap electricity without any pollution. This is because one gram of deuterium (heavy hydrogen) yields about 100,000 kWh of energy.

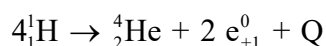
### 27.6.1 Energy in the Sun and Stars

The stars like our sun are very massive objects. They have been continuously emitting tremendous amount of energy for the last billions of years.

Such a huge amount of energy cannot be obtained by burning conventional fuels like coal. Nuclear fission can also not be the source of this energy, because heavy elements do not exist in the sun in large quantity. The sun mainly consists of hydrogen and helium gases. Then you may like to know: What is the source of energy in the sun? This question has engaged human intellect for long. As a child, you must have gazed the sky when you learnt the rhyme: Twinkle twinkle little star, How I wonder what you are!

You may know that the huge mass of the sun produces extremely strong gravitational field, which compresses its constituent gases by enormous pressure resulting in the rise of temperature to millions of kelvin at its centre. It has been estimated that the temperature at the centre of the sun is 20 million kelvin. At such high temperatures and pressures, gas molecules travel at high speeds and collide setting in thermonuclear reaction and resulting in the release of large amount of energy.

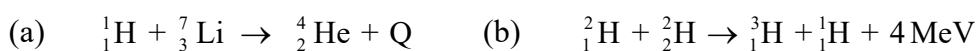
Hans Bethe proposed that fusion of hydrogen into helium is responsible for the energy produced in stars:



The overall result here is: four hydrogen nuclei fuse into a helium nucleus with the release of two positrons (electron-like microscope particles of the same mass but positive charge) and 26.8 MeV energy. ***The tremendous amount of energy released in a thermonuclear reaction is the source of energy in stars.*** The quantity of hydrogen in the sun is sufficient to keep it shining for nearly 8 billion years more.

### Intext Questions 27.6

- 200 MeV energy is released in fission of one  ${}_{92}^{235}\text{U}$  nucleus and 26.8 MeV energy is released in fusion of 4 protons. Which process releases more energy per unit mass?
- Consider the following reactions :



Calculate Q in the first reaction and mass of tritium in the second reaction. Given  $m({}_1^2\text{H}) = 2.014103\text{u}$ ,  $m({}_2^4\text{H}) = 4.002604\text{u}$ ,  $m({}_1^1\text{H}) = 1.007825\text{u}$  and  $m({}_3^7\text{Li}) = 7.015982\text{u}$ .

## 27.7 NUCLEAR ENERGY

We need energy for all activities in life. The amount of energy consumed per capita is a measure of advancement of a nation. According to a UNESCO report (2007), we are consuming about 40% more than what mother earth can generate in the form of food, water and energy. In fact, the human society has been continuously striving for energy security and looking for newer sources of energy. Due to over use, conventional sources of energy are depleting very fast and may exhaust completely in the next one hundred years. The nuclear energy is perhaps an important option for meeting our future energy needs through peaceful applications. Let us discuss these now.

### 27.7.1 Peaceful Applications

The most important peaceful application of nuclear energy is in the *generation of electricity*. One of the main advantages of nuclear power plant is that the fuel is not required to be fed into it continuously like the gas or coal in a thermal power plant. Further, it does not *pollute the environment to the extent discharge of smoke or ash from fossil fuel/power plants do*. The fuel once loaded in a reactor runs for nearly 6 months at a stretch. Because of this nuclear power plants have been used to power huge ships and submarines.

However, spent fuel from a reactor is highly radioactive because it contains a large number of radioisotopes. India has developed its own facility to treat this spent fuel and extract radioisotopes for use in agriculture, medicine, industry, and research. To avoid the spread of radioactive radiations from the radioactive wastes, the radioactive wastes are generally embedded deep inside salt mines in heavy steel cases. Yet, it has evoked considerable controversy due to its destructive potential which was displayed on August 6, 1945, when

an atom bomb was dropped on Hiroshima (Japan) killed hundred thousand people in a very short time. Subsequently, even more powerful hydrogen and nitrogen bombs have been developed. These can destroy this beautiful planet many times over.

### Nuclear Power in India

The possibility of harnessing nuclear power for civil use was recognised by Dr H.J. Bhabha soon after India got independence. He outlined a three stage development plan for meeting country's nuclear power needs. These are :

- Employ pressurised Heavy Water Reaction (PHWR) fuelled by natural uranium to generate electricity and produce plutonium as a by-product.
- Set up fast breeder reactors burning the plutonium to breed U-233 from thorium.
- Develop the second stage and produce a surplus of fissile material.

Nuclear power has been produced in India through 14 small and one mid-sized nuclear power reactors in commercial operation, eight under construction and more planned. As of now, nuclear power contributes nearly  $2 \times 10^{10}$  kW h of electricity – 3% of total power capacity available.

Government policy is to have 20 GW of nuclear capacity operating by 2020 and 25% nuclear contribution is foreseen by 2050.

### Intext Questions 27.7

1. What type of reactors are used in India for power generation?
2. How much  ${}_{92}^{235}\text{U}$  undergoes fission in an atomic bomb which releases energy equivalent to 20,000 tons of TNT. (Given that 1 g of TNT gives out 1000 calorie of heat).

### 27.7.2 Hazards of Nuclear Radiations and Safety Measures

The living and non-living things around us constitute our environment. In this environment, a delicate balance has existed for millions of years between the flora, fauna, aquatic and human life. This balance is now being threatened. One of the factors disturbing this balance is the ever increasing pollution in our environment. Out of the various types of pollutants present in our environment, the one which has very serious long term biological effects are the 'nuclear radiations'. Earlier these were present only because of natural sources like the radioactive minerals and cosmic rays, but now their presence is increasing day by day due to man-made sources. The major present day man-made sources of nuclear radiations are the nuclear tests, nuclear installations like the nuclear research facilities, nuclear reactors, and radio isotopes in treating diseases.

Nuclear radiations dissociate complex molecules of living tissues through ionization and kill the cells. They induce cancerous growth, cause sterility, severe skin burns, and lower the body resistance against diseases. They disrupt the genetic process, mainly in the unborn child, and show their effects even up to five generations. Nuclear radiations affect us not only directly, but also indirectly by affecting the flora, fauna and the aquatic life

around us. They kill vegetation, fishes and animals.

The damage caused by nuclear radiations depends on the exposed part of the body, as well as on the energy, intensity and the nature of the radiation. Different parts of human body show different sensitivities to radiation. The  $\alpha$ -particles are, as a rule, quite harmful because of their high ionizing power. The damaging effects of different radiations are generally compared in terms of their 'relative biological effectiveness', called the RBE factors. These factors for different particles/rays are given in Table-27.3.

**Table - 27.3 : RBE factors of different radiations**

Particle/rays	RBE factors
X-rays, $\gamma$ -rays, $\beta$ -particles	1
Thermal neutrons	2 to 5
Fast neutrons	10
$\alpha$ -particles, high energy ions of O, N, etc.	10 to 20

There is no control on natural sources of radiation. However, efforts can certainly be made to lower down radiation from man-made sources. Some of these are to:

- Avoid nuclear explosions.
- Minimize production of radio-isotopes.
- Extreme care should be exercised in the disposal of industrial wastes containing traces of radio-nuclides.
- Nuclear medicines and radiation therapy should be used only when absolutely necessary, and with well-considered doses.

## WHAT YOU HAVE LEARNT

- The nucleus in an atom contains positively charged protons and uncharged neutrons.
- The number of protons inside the nucleus of an atom of any element gives the atomic number of the element.
- The sum of the number of protons and neutrons in the nucleus of an atom is called its mass number.
- The atoms having same atomic number but different mass numbers are called isotopes.
- The atoms with same mass number but different atomic numbers are called isobars.
- The atoms with same number of neutrons are called isotones.
- The nucleons inside the nucleus of every atom are bound together by strong attractive nuclear forces which are short-range and charge-independent.
- The mass of a nucleus is found to be less than the sum of the masses of its nucleons. This difference in mass is called mass-defect. It is a measure of the binding energy.
- The size (volume) of the nucleus depends on its mass number.

- The spontaneous emission of  $\alpha$ -particles or  $\beta$ -particles followed by  $\gamma$ -emission from any nucleus is called radioactivity.
- The  $\alpha$ -particles have been identified as helium nuclei, while  $\beta$ -particles have been identified as fast moving electrons. The  $\gamma$ -rays are electromagnetic waves of extremely short wavelength.
- According to the law of radioactive decay, the number of radioactive atoms disintegrating per second is proportional to the number of radioactive atoms present at that instant.
- The half-life of a radioactive substance is the time during which the number of radioactive atoms reduces to half of its original number.
- The law of exponential decay is  $N(t) = N_0 \exp(-\lambda t)$ .
- Valence electrons take part in chemical reactions and the energy involved in such reactions is of the order of 1 eV.
- In a nuclear reaction, the atomic nuclei interact to form a new element.
- Energy involved in nuclear reaction is of the order of MeV.
- In a nuclear reaction, atomic number, mass number and charge are conserved.
- When a heavy nucleus like uranium is bombarded by slow neutrons, it splits into two fragments with release of 2-3 neutrons and 200 MeV energy. This process is known as nuclear fission.
- Substances that undergo fission are called fissile substances.  $^{233}\text{Th}$ ,  $^{233}\text{U}$ ,  $^{235}\text{U}$ , and  $^{239}\text{Pu}$  are fissile materials.
- Chain reaction occurs when more than one emitted neutron induces further fission for each primary fission.
- Nuclear reactor is a device to sustain controlled chain reaction.
- In nuclear fusion two light nuclei are fused into one.

### TERMINAL EXERCISE

1. When does a radioactive sample disintegrate?
2. Differentiate between isotopes and isobars.
3. Explain the characteristics of binding energy per nucleon versus mass number curve.
4. What is the nature of nuclear force? Give its characteristics.
5. Explain how decay constant is related to half-life of a radioactive substance.
6. Define the following terms:
 

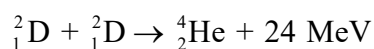
(i) Atomic number;	(ii) Mass number;	(iii) Mass defect;
(iv) Binding energy of nucleons;	(v) Half-life	(vi) Average life;
(vii) Decay constant.		



7. State the law of radioactive decay.
8. What is carbon dating? What is its importance?
9. Calculate the number of neutrons, protons, and electrons in the following atoms.
  - (i)  ${}_{11}^{23}\text{Na}$ ;    (ii)  ${}_{1}^2\text{H}$ ;    (iii)  ${}_{92}^{238}\text{U}$ ;    (iv)  ${}_{17}^{35}\text{Cl}$
10. Calculate the mass defect and binding energy of nucleons for the following nuclei.
  - (i)  ${}_{2}^4\text{He}$ ;    (ii)  ${}_{3}^7\text{Li}$ ;    (iii)  ${}_{7}^{14}\text{N}$ ;

Given,  $1\text{ u} = 1.660566 \times 10^{-27}\text{ kg} = 931\text{ MeV}$ , mass of a proton =  $1.007276\text{ u}$ , mass of a neutron =  $1.008665\text{ u}$ , mass of  ${}_{2}^4\text{He}$  atom =  $4.00260\text{ u}$ , mass of  ${}_{3}^7\text{Li}$  atom =  $7.01601\text{ u}$ , mass of  ${}_{7}^{14}\text{N}$  atom =  $14.00307\text{ u}$ .

11. Using the present-day abundance of the two main uranium isotopes and assuming that the abundance ratio could never have been greater than unity, estimate the maximum possible age of the earth's crust. (Given that the present-day ratio of  ${}^{238}\text{U}$  and  ${}^{235}\text{U}$  is 1378: 1; Half-life of  ${}^{238}\text{U}$  is  $4.5 \times 10^9$  year; and that of  ${}^{235}\text{U}$  is  $7.13 \times 10^8$  years.)
12. If the activity of a radioactive sample drops to  $1/16^{\text{th}}$  of its initial value in 1 hour and 20 minutes. Calculate the half-life.
13. How does a nuclear reaction differ from a chemical reaction?
14. What is the use of moderator and absorber in a fission reactor?
15. Based on B.E per nucleon versus mass number curve, explain nuclear fusion.
16. What is a nuclear reaction? State the conservation laws obeyed in nuclear reactions. Give three examples of nuclear reactions.
17. What is nuclear fission? Give an example to illustrate your answer.
18. Calculate the mass of  ${}^{235}\text{U}$  consumed to generate 100 megawatts of power for 30 days.
19. Heavy hydrogen undergoes the following fusion reaction.



Calculate the amount of heavy hydrogen used in producing the same energy as above. Compare the two results.

20. What is nuclear fusion? Write an equation of nuclear fusion to support your answer.
21. What is the source of energy in the sun? How is it generated? Illustrate with an example.
22. Describe the construction of an atomic reactor.
23. Calculate the energy released in a fusion reaction,  $3({}_{2}^4\text{He}) \rightarrow {}_{6}^{12}\text{C}$   
given, the mass of an  $\alpha$ -particle =  $4.00263\text{u}$ .

## ANSWERS TO INTEXT QUESTIONS

## 27.1

1.	Isotopes	Isobars	Isotones
	${}^1_6\text{C} \text{ \& } {}^{14}_6\text{C}$	${}^{76}_{32}\text{Ge} \text{ \& } {}^{76}_{34}\text{Se}$	${}^2_1\text{H} \text{ \& } {}^3_2\text{He}$
	${}^1_1\text{H} \text{ \& } {}^2_1\text{H} \text{ \& } {}^3_1\text{H}$	${}^{40}_{18}\text{A} \text{ \& } {}^{40}_{20}\text{Ca}$	${}^{14}_6\text{C} \text{ \& } {}^{18}_8\text{O}$
	${}^{16}_8\text{O} \text{ \& } {}^{18}_8\text{O}$	${}^{76}_{32}\text{Ge} \text{ \& } {}^{76}_{34}\text{Se}$	${}^{23}_{11}\text{Na} \text{ \& } {}^{24}_{12}\text{Mg}$
	${}^{35}_{17}\text{Cl} \text{ \& } {}^{37}_{17}\text{Cl}$	${}^3_1\text{H} \text{ \& } {}^3_2\text{He}$	${}^{27}_{13}\text{Al} \text{ \& } {}^{28}_{14}\text{Si}$
	${}^{206}_{82}\text{Pb} \text{ \& } {}^{207}_{82}\text{Pb}$	${}^7_3\text{Li} \text{ \& } {}^7_4\text{Be}$	${}^{27}_{13}\text{Al} \text{ \& } {}^{28}_{14}\text{Si}$
	${}^{238}_{92}\text{U} \text{ \& } {}^{239}_{92}\text{U}$		

2. Heavier                                    3. Mass                                    4. Nucleons  
 5. 14, 14                                    6. Atomic                                    7. (ii) Atomic number.

## 27.2

1.  $\Delta m = 1.041358 \text{ u}; 969.5 \text{ MeV}$ .  
 2.  $2.4 \times 10^{-15} \text{ m}$ .

## 27.3

1. Nuclear disintegration usually involves  $\alpha$  or  $\beta$  emission which results in change of atomic and mass numbers of the parent element. With the emission of  $\alpha$  and  $\beta$  particles, the heavier nuclei shed some of their mass resulting in comparatively lighter nuclei. Hence, it is a nuclear disintegration phenomenon.
2. Ionizing power of  $\alpha > \beta > \gamma$   
 Penetration power of  $\alpha < \beta < \gamma$
3. (i)  $a = Z - 2$  and  $b = A - 4$   
 (ii)  $a = Z + 1$  and  $b = A$ .
4. Two half-life times are required – one for reduction from 10 to 5 grams and the other from 5 to 2.5 grams, i.e., 10 years.

## 27.4

1. a)  ${}^1_9\text{F} + {}^1_1\text{H} \rightarrow {}^1_8\text{O} + {}^4_2\text{He}$ ;      b)  ${}^{27}_{13}\text{Al} + {}^1_0\text{n} \rightarrow {}^{24}_{11}\text{Na} + {}^4_2\text{He}$ ;  
     c)  ${}^{234}_{90}\text{Th} \rightarrow {}^{234}_{90}\text{Pa} + {}^0_{-1}\text{e}$ ;      d)  ${}^{63}_{29}\text{Cu} + {}^2_1\text{D} \rightarrow {}^{64}_{30}\text{Zn} + {}^1_0\text{n}$
2. 17.9 MeV
3.  ${}^{14}_7\text{N} + {}^4_2\text{He} \rightarrow {}^{17}_8\text{O} + {}^1_1\text{H} + 6.5 \text{ MeV}$ .

## 27.5

1. Due to increase of  $n/p$  ratio above the natural ratio, its stability decreases. To decrease the ratio to attain more stability, it emits a  $\beta$ -particle.
2. 239 Pu
3. 200 MeV.

## 27.6

1. In fission the energy released is 0.84 MeV/u whereas in fusion, it is 6.7 MeV/u. Thus energy released per unit mass is more in the latter case.
2. (a) 17.3 MeV,      (b) 2.69 MeV.

## 27.7

1. Pressurized Heavy Water Reactor
2. Nearly 1 kg.

## ANSWERS TO TERMINAL EXERCISE

9. (i) 12, 11, 11      (ii) 1, 1, 1      (iii) 146, 92, 921      (iv) 18, 17, 17
10. (i) 0.034, 28MeV      (ii) 0.044, 37.86 MeV      (iii) 0.10854, 101MeV
11.  $6.0 \times 10^9$  years
12. 20 min
18. 30.6 kg
19. 146.6 g
23. 7.35 MeV



# SEMICONDUCTOR DEVICES

## INTRODUCTION

In our day-to-day life, transistor radio, TV, cell phone and computers what we use are called semiconductor devices. Silicon and germanium are the most familiar semiconductor materials. Normally, the conductivity of a semiconductor lies in-between the conductivities of metals and insulators. It should be noted here that at absolute zero, the semiconductor also acts like a perfect insulator. The conductivity of a semiconductor is influenced by adding some impurity element called dopant. Depending on the type of carrier added by a dopant, the semiconductor is classified as p-type or n-type.

If one side of a pure semiconductor is doped with p-type impurity and the other side is doped with n-type impurity, then we call a p-n junction. A p-n junction is also called a diode. In this chapter you will learn about various types of semiconductors, their behavior and how they are combined to form useful devices such as solar cell, photodiode, transistor, etc. These simple structures are used in voltage regulators, display switches, storage devices, communication systems, computers, satellites, space vehicles and power systems. At the end of this chapter logic gates were discussed. Logic gates are electronic devices which act as building blocks for digital circuits.

### OBJECTIVES

After studying this lesson, you should be able to

- *differentiate between n-type and p-type semiconductors;*
- *explain formation of depletion region and barrier potential in a p-n junction diode;*
- *describe I-V characteristics of a p-n junction diode in the forward and reverse biases;*
- *know the applications of diodes and triodes;*
- *explain the action of a transistor; their input and output characteristics;*
- *explain how Logic gates follow Boolean algebra?*

### 28.1 TYPES OF MATERIALS

Before discussing the classification of semiconductors, it is necessary to understand how materials are classified? On the basis of resistivity values, materials are classified into three categories, namely metals, semiconductors and insulators.

- (i) **Metals** : They possess very low resistivity (or high conductivity). The values are:  
 Resistivity ( $\rho$ ) for metals =  $10^{-2} - 10^{-8} \Omega\text{m}$   
 Conductivity ( $\sigma = 1/\rho$ ) =  $10^2 - 10^8 \text{ Sm}^{-1}$
- (ii) **Semiconductors** : The resistivity or conductivity of semiconductors lie between metals and insulators  
 Resistivity ( $\rho$ ) for semiconductors =  $10^{-5} - 10^6 \Omega\text{m}$   
 Conductivity ( $\sigma$ ) =  $10^5 - 10^{-6} \text{ Sm}^{-1}$
- (iii) **Insulators** : Insulators have high resistivity (or low conductivity)  
 Resistivity ( $\rho$ ) for insulators =  $10^{11} - 10^{19} \Omega\text{m}$   
 Conductivity ( $\sigma$ ) =  $10^{-11} - 10^{-19} \text{ Sm}^{-1}$

The values mentioned above are approximated values and the exact criteria for the classification of materials are explained by the energy-band theory concept.

### Energy-band theory

An isolated atom possesses discrete energies for its different electrons. Suppose two isolated atoms are brought into very close proximity; then, the electrons in the orbits of these two atoms interact with each other. As a result, in the combined system, the energies of electrons will no longer remain at the same level; they will change slightly, becoming both lower and higher than their original values. Consequently, at each energy level, two closely spaced energy levels exist.

If 'N' number of atoms are brought together to form a solid and their electrons interact, a set of 'N' closely spaced energy levels are formed, replacing the discrete energy levels. These are known as bands of allowed energies. Between the bands of allowed energies, there are empty energy regions called the forbidden bands. There are mainly three energy bands in solids and these energy bands are important to understand the behavior of solids.

The three energy bands are:

- Valence band
  - Conduction band
  - Forbidden band or forbidden gap
- **Valence band** : The energy band which is formed by grouping the range of energy levels of the valence electrons or outermost orbit electrons is called as the valence band. Valence band is present below the conduction band as shown in Fig. 28.1. Electrons in the valence band have lower energy than the electrons in conduction band. The electrons present in the valence band are loosely bounded to the nucleus of an atom.
- **Conduction band** : The energy band which is formed by grouping the range of energy levels of the free electrons is called as the conduction band. Generally, the conduction band is empty but when external energy is applied the electrons in the valence band jumps into the conduction band and become free electrons. Electrons

in the conduction band have higher energy than the electrons in the valence band. The conduction band electrons are not bound to the nucleus of the atom.

- **Forbidden band/gap :** The energy gap which is present between the valence band and conduction band by separating the two energy bands is called as forbidden band or forbidden gap. In solids, electrons cannot stay in forbidden gap because there is no allowed energy state in this region. Forbidden gap is the major factor for determining the electrical conductivity of a solid. The classification of materials as insulators, conductors and semiconductors depends on the forbidden gap.

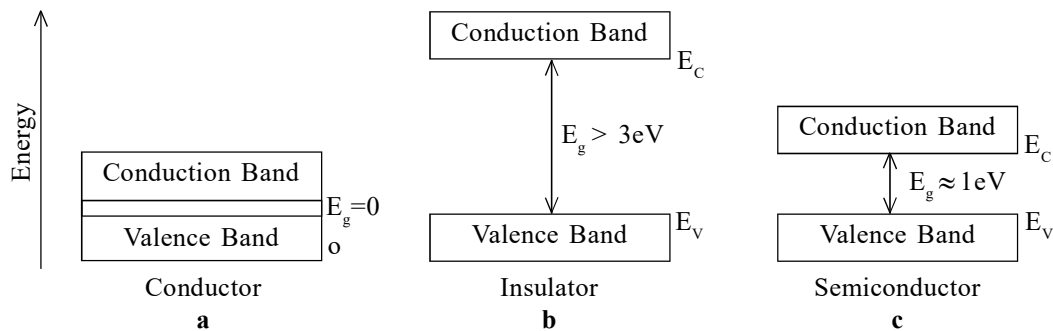
The energy associated with the forbidden band is called the energy gap, and it is measured in units of electron volts (eV).

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

The applied external energy in the form of heat or light must be equal to the forbidden gap in order to push an electron from valence band to the conduction band.

### Classification of materials based on band theory of solids

The forbidden gap plays a significant role in determining the electrical conductivity of a material. Based on the forbidden gap, materials are also classified into three types, as shown in the Fig. 28.1.



**Fig. 28.1 :** (a-c) Energy band diagrams for metals (conductors), insulators and semiconductors.

- **Metals :** The energy band diagram of metals is shown in Fig. 28.1 (a). Here, the conduction band and valence bands are partially filled and partially empty. Therefore, conduction and valence bands get overlapped, and hence electrons from the valence band can easily move into the conduction band. This situation makes a large number of electrons available for electrical conduction. Here it should be noted that electrons from the lower level can move to the higher level and making conduction possible. Finally, the resistance of such materials is low or conductivity is high.
- **Insulators :** The energy band diagram of insulators is shown in Fig. 28.1 (b). A large band gap ( $E_g$ ) exists and there are no electrons in the conduction band. Therefore, no electrical conduction is possible in this case. Here it should be noted that the energy gap is so large ( $E_g > 3 \text{ eV}$ ) and the electrons cannot be excited from the valence band to the conduction band by any thermal excitation.
- **Semiconductors :** The energy band diagram of semiconductors is shown in Fig. 28.1 (c). Here, a finite small band gap ( $E_g$ ) exists. Because of the small gap,

even at room temperature some electrons from the valance band acquire some energy ( $E_g \approx 1 \text{ eV}$ ) and move to the conduction band. Though electrons are small in number, they can move in the conduction band. Here a noteworthy aspect is that the resistance of semiconductors is not as high as that of insulators.

### Intext Questions 28.1

1. On the basis of \_\_\_\_\_ values, materials are classified into the three categories, namely metals, insulator and semiconductors.
2. Conductivity is reciprocal to \_\_\_\_\_
3. Electrons in the conduction band has \_\_\_\_\_ energy, compared to electrons of the valance band.
4. The unit of energy gap is \_\_\_\_\_
5. The energy gap of insulators is generally \_\_\_\_\_ than 3 eV.

## 28.2 CLASSIFICATION OF SEMICONDUCTORS

Based on the purity, semiconductors are classified as intrinsic (pure) and extrinsic (impure) semiconductors. A detailed aspect of this is as follows:

### 28.2.1 An Intrinsic Semiconductor

Pure silicon and germanium are intrinsic semiconductors. They have no impurity and the electrons in these elements are all tightly locked in their crystalline structure. Therefore, electrons cannot move freely. When energy is added to pure silicon in the form of heat, it can cause a few electrons to break free of their bonds, leaving behind a hole in each case. The absence of electrons is treated as a positively charged particle (hole). It is having the same amount of charge as an electron, but the charge is positive. These electrons move randomly in the crystal. These electrons and holes are called free carriers and can move to create electrical current. However, they are few, so may not very useful in pure silicon. Since there are no excess electrons or holes, in this case, the number of electrons and holes present at any given time will always be equal.

It should be noted here that in an intrinsic semiconductor, electrons and holes are always generated in pairs and the negative charge of free electrons is exactly balanced by the positive charge of holes. However, a hole only shifts its position due to the motion of an electron from one place to another. So, we can say that when a free electron moves in a crystal because of thermal energy; its path deviates whenever it collides with a nucleus or other free electrons. This gives rise to a zig-zag or random motion, which is similar to that of a molecule in a gas.

Consider the electron- hole pair [Fig. 28.2 (a)] is generated at point A. The free electron drifts in the crystal leaving behind a hole. The broken bond now has only one electron and this unpaired electron has tendency to acquire an electron and complete its pair by forming a covalent bond. Due to thermal energy, the electron from neighboring bond, say at point B, may get excited to break its own bond and jump into the hole at A. As a result, the hole

at A vanishes and a new hole appears at B [Fig. 28.2 (c)]. Thus, the motion of electron from point B to point A causes the hole to move from A to B. What happens when hole at B attracts and captures a valence electron from neighbouring bond at C? The movement of electron from C to B causes movement of hole from B to C [see Fig. 28.2 (d) and (e)]. Conventionally, the flow of electric current through the semiconductor is taken in the same direction in which holes move.

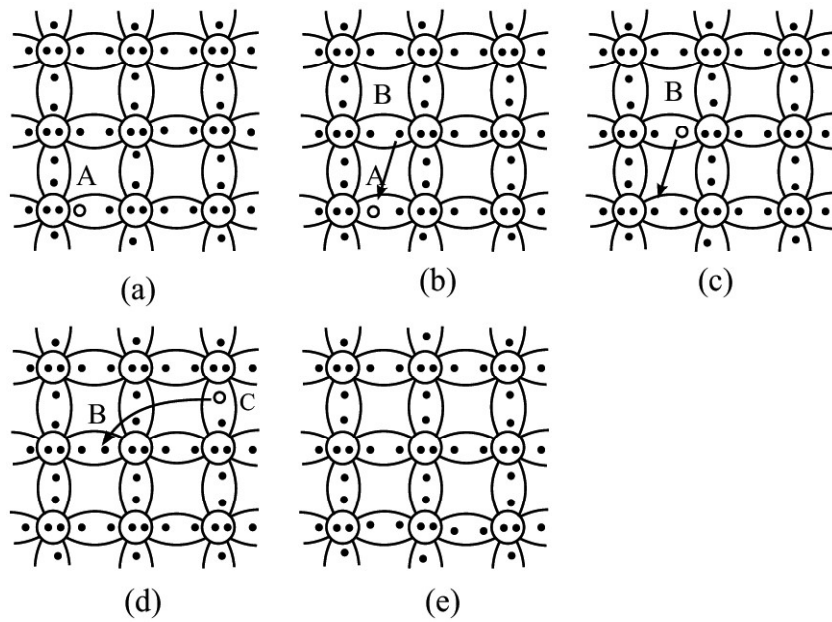


Fig. 28.2 : Movement of electrons and holes in a semiconductor

At absolute zero temperature, all valence electrons are tightly bound to their parent atoms and intrinsic semiconductor behaves as an insulator. At room temperature, the thermal energy makes a valence electron in an atom to move away from the influence of its nucleus. Therefore, a covalent bond is broken and electron becomes free to move in the crystal, resulting in the formation of a vacancy, called hole. **Thus, due to thermal energy, some electron-hole pairs are generated, and semiconductor exhibits small conductivity.** For example, at room temperature (300 K), Ge has intrinsic carrier concentration of about  $2.5 \times 10^{19} \text{ m}^{-3}$ . As temperature increases, more electron-hole pairs are generated and conductivity increases. Alternatively, we can say that resistivity decreases as temperature increases. It means that semiconductors have negative temperature coefficient of resistance.

### 28.2.2 An Extrinsic Semiconductor

As mentioned earlier, intrinsic semiconductors have high resistivity, and their conductivity shows little flexibility. For these reasons, intrinsic (pure) semiconductors are of little use; at best, they can be used as heat or light-sensitive resistors. These limitations are overcome by adding a small and measured quantity of another material to intrinsic (pure) semiconductor, which either increases the number of holes or electrons.

The process of addition of impurities to a pure or intrinsic semiconductor is called **doping** and the impurity atom that is added is called **dopant**. Such doped semiconductors are called **extrinsic** semiconductors. The dopants are generally taken from either **group III** (having three valence

	III	IV	V	VI
II	Al	Si	P	S
Zn	Ga	Ge	As	Se
Cd	In	Sn	Sb	Te
Hg				

Fig. 28.3 : A part of the Periodic Table. Group III and V elements are used for doping an intrinsic semiconductor.



electrons) or **group V** (having five valence electrons) of the Periodic Table. Fig. 28.3 shows a small portion of the Periodic Table. Here groups III and V have been highlighted to indicate the types of materials generally used for doping.

Normally we add a very small amount of impurity atoms to the pure semiconductor. It is of the order of one impurity atom per  $10^8$  atoms of intrinsic semiconductor. These atoms change the balance of charge carriers; either they add free electrons or create holes. Either of these additions makes the material more conducting. Thus, most of the charge carriers in extrinsic semiconductors originate from the impurity atoms.

### 28.2.3 n-and p-type Semiconductors

From the electronic configuration of Si ( $1s^2, 2s^2, 2p^6, 3s^2, 3p^2$ ), you will recall that ten electrons are tightly bound to the nucleus and four electrons revolve around the nucleus in the outermost orbit. In an intrinsic silicon semiconductor, the Si atom attains stability by sharing one electron each with four neighboring Si atoms. (This is called covalent bonding). The same holds true for germanium; its electronic configuration is  $1s^2, 2s^2, 2p^6, 3s^2, 3p^6, 3d^{10}, 4s^2, 4p^2$ . When silicon (or germanium) is doped with a pentavalent (five electrons in the outermost orbit) atom like phosphorus, arsenic or antimony, four electrons form covalent bonds with the four neighboring silicon atoms, but the fifth (valence) electron remains unbound and is available for conduction, as shown in Fig. 28.4 (a). Thus, when a silicon (or germanium) crystal is doped with a pentavalent element, it develops excess free electrons and is said to be an n-type semiconductor. Such impurities are known as donor impurities. Note that in n-type semiconductors, the number of free electrons is far greater than the number of holes and the letter (n) stands for negative charges.

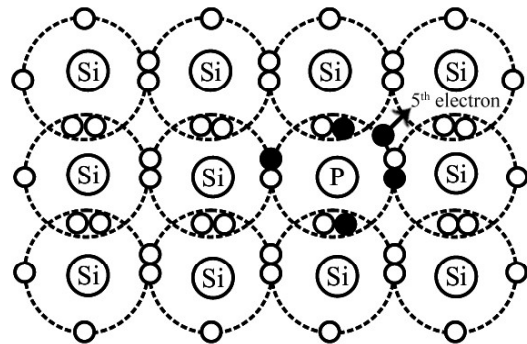


Fig. 28.4 (a) : Covalent bonding in a n-type semiconductor

If silicon (or germanium) is doped with a trivalent (three electrons in the outermost shell) atom like boron, aluminum, gallium or indium, three valence electrons form covalent bonds with three silicon atoms and deficiency of one electron is created. This deficient of electron is referred to as hole. It is shown in Fig. 28.4 (b). Such a semiconductor is said to be a p-type semiconductor and the impurities are known as acceptor impurities.

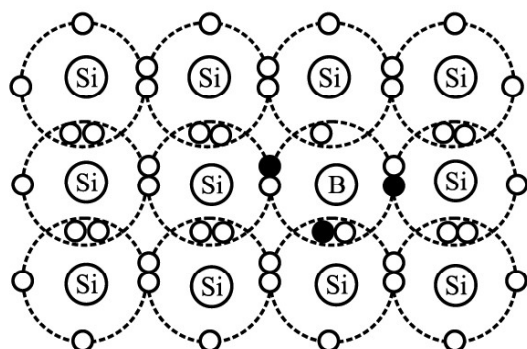


Fig. 28.4 (b) : Covalent bonding in a p-type semiconductor

Now, a fundamental question arises: Is an n-type semiconductor negatively charged? The answer to this question is not in the affirmative. In fact, the number of free electrons is exactly equal to the total number of holes and positively charged ions in a semiconductor, whether intrinsic or doped. Therefore, a semiconductor is electrically neutral.

Note that in a p-type semiconductor, more holes are created due to addition of acceptor impurity than by breaking covalent bonds due to thermal energy at room temperature. Hence, the net concentration of holes is significantly greater than that of electrons. That is, in a p-type semiconductor, the holes are the majority charge carriers.

Similarly in n-type semiconductors more electrons are created due to the addition of donor type impurity atoms. Hence, concentration of electrons are much greater than that of electrons which are normally generated due to thermal energy.

### Intext Questions 28.2

- At 300 K, pure silicon has intrinsic carrier concentration of  $1.5 \times 10^{16} \text{ m}^{-3}$ . What is the concentration of holes and electrons?
- The n-type semiconductor is obtained by doping with
  - trivalent impurity
  - pentavalent impurity
  - tetravalent impurity
  - trivalent as well as tetravalent
- An intrinsic semiconductor can be converted into an extrinsic semiconductor by addition of \_\_\_\_\_. This process is called \_\_\_\_\_.
- Electrons in n-type semiconductor and holes in p-type semiconductor are the \_\_\_\_\_ carriers.
- An extrinsic semiconductor has \_\_\_\_\_ resistivity as compared to an intrinsic semiconductor.

### 28.3 p-n JUNCTION

In the last section, we have studied n-type and p-type semiconductors, which respectively have electrons and holes as majority charge carriers. What do you think if an n-type material is placed in contact with a p-type material? The answer to this question will be discussed in the following section.

#### 28.3.1 Formation of a p-n Junction

To form a p-n junction, the most convenient way is to introduce donor impurities on one side and acceptor impurities into the other side of a single semiconducting crystal, as shown in Fig. 28.5.

We know that, there is a greater concentration of electrons in the n-region of the crystal and of holes in the p-region. Because of this, electrons tend to diffuse to the p-region and holes to the n-region and finally get recombined. Each recombination eliminates a hole and a free electron. This results in creation of positively and negatively charged ions near the junction

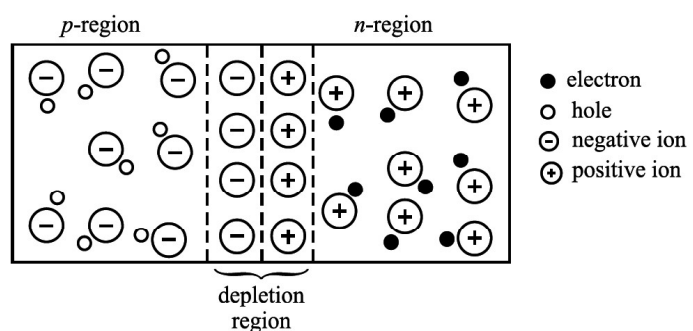


Fig. 28.5 : A p-n junction with depletion region

in n and p regions, respectively. As these charges accumulate, they tend to act as shield preventing further movement of electrons and holes across the junction. After the process of recombination, a narrow region near the junction is depleted in mobile charge carriers (about 0.5 mm thick), called the depletion region or space-charge region. Due to accumulation of charges near the junction, an electric field is established. This gives rise to electrostatic potential, known as barrier potential. This barrier has polarities, as shown in Fig. 28.6. When there is no external electric field, this barrier prevents diffusion of charge carriers across the junction.

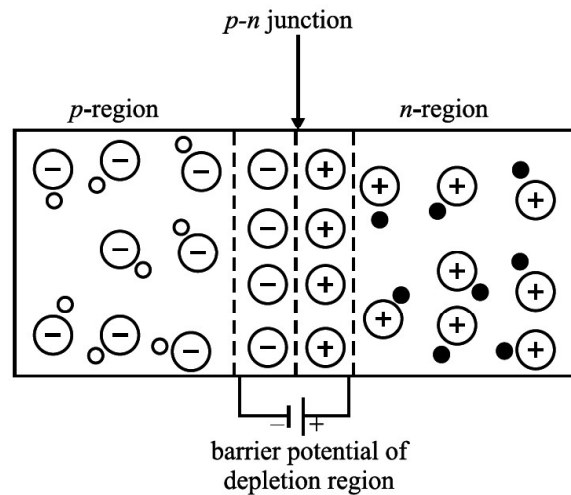
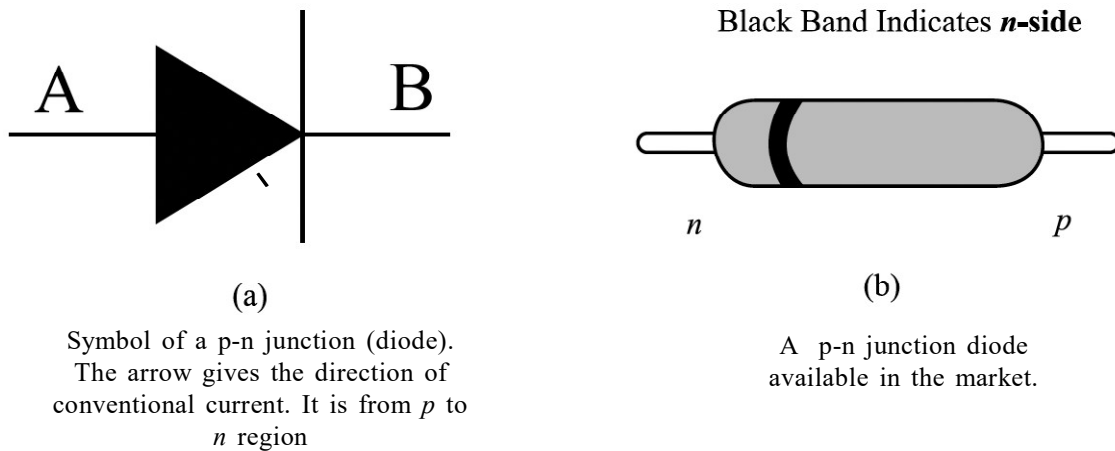


Fig. 28.6 : Barrier potential due to depletion region

The barrier potential is characteristic of the semiconductor material. It is about 0.3 V for Ge and about 0.7 V for Si. The junction acts as a diode. It is symbolically represented as shown in Fig. 28.7 (a). Here A corresponds to p-region and acts as an anode. Similarly, B indicates n-region and corresponds to a cathode. Fig. 28.7 (b) shows a picture of p-n junction diode available in market.



(a) Symbol of a p-n junction (diode). The arrow gives the direction of conventional current. It is from p to n region

(b) A p-n junction diode available in the market.

Fig. 28.7 : p-n junction diode

You might have noted that semiconductor diodes are designated by two letters followed by a serial number. The first letter indicates the material: **A** is used for material with a band gap of 0.6 eV to 1.0 eV such as germanium. **B** is used for material with a band gap of 1.0 eV to 1.3 eV, such as silicon. The second letter indicates the main application: **A** signifies detection diode, **B** denotes a variable capacitance diode, **E** for tunnel diode, **Y** for rectifying diode and **Z** denotes Zener diode. The serial numbers specify power rating, peak reverse voltage, maximum current rating, etc. (We have to refer to manufacturer’s catalogue to know exact details.) For example, BY127 denotes a silicon rectifier diode and BZ148, represents a silicon Zener diode.

To make visual identification of anode and cathode, the manufacturers employ one of the following ways:

- the symbol is painted on the body of the diode;
- red and blue marks are used on the body of the diode. Red mark denotes anode, whereas blue indicates the cathode;
- a small ring is printed at one end of the body of the diode that corresponds to the cathode. The band in Fig. 28.7 (b) indicates the n-side of the p-n junction.

### Intext Questions 28.3

1. Fill in the blanks:
  - (a) When a p-n junction is formed, the \_\_\_\_\_ diffuse across the junction.
  - (b) The region containing uncompensated acceptor and donor ions is called \_\_\_\_\_ region.
  - (c) The barrier potential in silicon is \_\_\_\_\_ V and in germanium, it is \_\_\_\_\_ V.
  - (d) In a p-n junction with no applied electric field, the electrons diffuse from n-region to p-type region as there is \_\_\_\_\_ concentration of \_\_\_\_\_ in n-region as compared to p-region.
2. Choose the correct option:
  - (a) The potential barrier at the p-n junction is due to the charges on the either side of the junction. These charges are
    - (i) majority carriers
    - (ii) minority carriers
    - (iii) fixed donor and acceptor ions.
    - (iv) none of above
  - (b) In a p-n junction without any external voltage, the Junction current at equilibrium is,
    - (i) due to diffusion of minority carriers only
    - (ii) due to diffusion of majority carriers only
    - (iii) zero, as no charges are crossing the junction
    - (iv) zero, as equal and opposite charges are crossing the junction
  - (c) In a semiconductor diode, the barrier potential repels
    - (i) minority carriers in both the regions
    - (ii) majority carriers in both the regions
    - (iii) both the majority and the minority carriers
    - (iv) none of the above
3. Why depletion region is named so? What is depletion region made of?

## 28.4) FORWARD AND REVERSE BIAS OF p-n JUNCTION

The term biasing means application of voltage. To make a p-n junction to conduct, electrons should move from the n-type region to the p-type region and holes moving in the reverse direction. To overcome the potential barrier across the junction, a battery should be connected to the two ends of the p-n junction diode. The battery can be connected to the p-n junction in two ways:

1. Positive terminal of the battery connected to the p-side and negative terminal of the battery connected to the n-side. This is called **forward bias** [Fig. 28.8 (a)].
2. Positive terminal of the battery connected to the n-side and negative terminal of the battery connected to the p-side. This is called **reverse bias** [Fig. 28.8 (b)].

When a junction is forward biased and the bias exceeds barrier potential, holes are compelled to move towards the junction and cross it from the p-region to the n-region. Similarly, electrons cross the junction in the reverse direction. This sets in **forward current** in the diode. The current increases with voltage and is of the order of a few milliamperes (mA). Under the forward bias condition, the junction offers low resistance to flow of current. The value of junction resistance, called **forward resistance**, is in the range  $10\ \Omega$  to  $30\ \Omega$ .

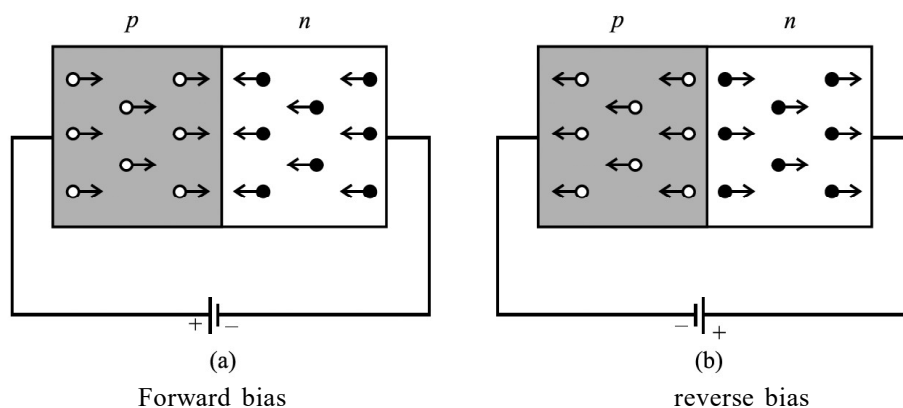


Fig. 28.8 : Biasing condition of p-n junction

If p-n junction is **reverse biased**, then the holes in the p-region and electrons in the n-region move away from the junction. Does it mean that no current shall flow in the circuit? No, a small current does flow even now because of the fewer number of electron-hole pairs generated due to thermal excitations. This small current caused by minority carriers is called **reverse saturation current** or **leakage current**. In most of the commercially available diodes, the reverse current is almost constant and independent of the applied reverse bias. Its magnitude is of the order of a few microamperes ( $\mu\text{A}$ ) for Ge diodes and nano-amperes (nA) in Si diodes.

A p-n junction offers low resistance when forward biased, and high resistance when reverse biased. This property of p-n junction is used for ac-rectification.

When the reverse bias voltage is of the order of a few hundred volts, the current through the p-n junction increases rapidly and damages it due to excessive power dissipation. The voltage at which a diode breaks down is termed as **breakdown voltage**. Physically, it can

be explained as follows: When a reverse bias is applied, a large electric field is established across the junction. This field (i) accelerates the available minority carriers, which, in turn, collide with the atoms of the semiconductor material and eject more electrons through energy transfer (avalanche effect), and (ii) breaks covalent bonds by exerting large force on electrons bound by the bonds. This results in creation of additional electron-hole pairs in the junction region (Zener effect). Both these processes give rise to large reverse current even for a small increment in reverse bias voltage. This process is termed as Zener breakdown.

### Intext Questions 28.4

1. What is forward bias?
2. What is reverse bias?
3. Fill in the blanks:
  - (a) When forward bias is applied on a p-n junction diode, the width of the depletion region \_\_\_\_\_.
  - (b) When a p-n junction diode is reverse biased, the width of depletion region \_\_\_\_.
  - (c) When the reverse bias voltage is made too high, the current through the p-n junction \_\_\_\_\_ abruptly. This voltage is called \_\_\_\_\_.
4. Choose the correct option:
  - (a) In a forward biased junction
    - (i) the holes in the n-region move towards the p-region
    - (ii) there is movement of minority carriers
    - (iii) charge carriers do not move
    - (iv) majority carriers in both the regions (n and p-regions) move into other regions
  - (b) In a reverse biased junction
    - (i) there is no of potential barrier
    - (ii) there is movement of majority carriers only
    - (iii) there is movement of minority carriers only
    - (iv) none of the above
5. State two types of reverse breakdowns which can occur in a p-n junction diode.

### 28.5 CHARACTERISTICS OF p-n JUNCTION DIODE

The practical application of a semiconductor device in electronic circuits depends on the current and voltage (I-V) relationship, as it gives vital information to a circuit designer as well as a technician. Therefore, with the help of I-V characteristics, one can know how much current flows through the junction diode at a particular voltage.

### 28.5.1 Forward Bias Characteristics

Fig. 28.9 (a) shows the forward bias characteristic of a p-n junction diode, the positive terminal of a battery (B) is connected to p-side of the diode through the rheostat (alternatively, we can use a variable battery). The voltage applied to the diode can be varied with the help of the rheostat. The milliammeter (mA) measures the current in the circuit and voltmeter (V) measures the voltage across the diode. The direction of conventional current is the same as the direction of the diode arrow. Since current experiences little opposition to its flow through a forward biased diode and it increases rapidly as the voltage is increased, a resistance (R) is added in the circuit to limit the value of current. If this resistance is not included, the diode may get permanently damaged due to flow of excessive current through it. I-V characteristic curve of a p-n junction in forward bias is shown in Fig. 28.9 (b).

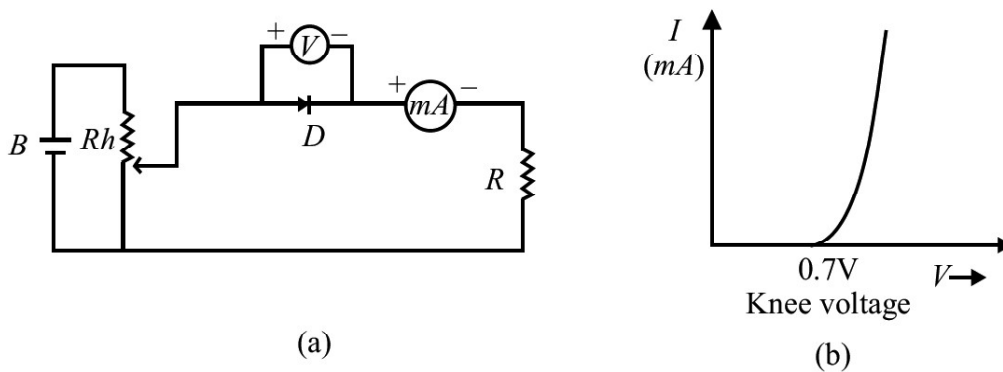


Fig. 28.9 : (a) Circuit diagram I-V characteristics of a p-n junction diode in forward bias  
(b) Typical characteristics curve

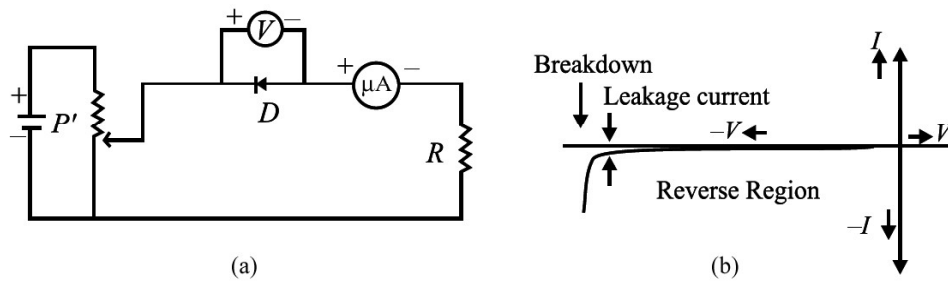
Here one should note that the characteristic curve does not pass through origin; instead, it meets the V-axis around 0.7 V. It means that the p-n junction does not conduct until a definite external voltage is applied to overcome the barrier potential. The forward voltage required to get the junction in conduction mode is called **knee voltage**. It is about 0.7 V for Si and 0.3 V for Ge p-n junction.

This voltage is needed to start the hole-electron combination process at the junction. If the applied voltage is increased beyond knee voltage, then the current through the diode increases linearly. For voltage of around 1 V (one volt), the current may attain a value of 30-80 mA.

### 28.5.2 Reverse Bias Characteristics

To draw reverse bias characteristics of a p-n junction, we use the circuit diagram shown in Fig. 28.10 (a). A typical I-V characteristic curve of a p-n junction in reverse bias is shown in Fig. 28.10 (b). If you compare it with Fig. 28.9 (a) for forward I-V characteristics, you will note two changes:

- (i) The terminals of the junction are reversed.
- (ii) Instead of milliammeter, micro-ammeter ( $\mu\text{A}$ ) is used.



**Fig. 28.10** : (a) Circuit diagram to obtain I-V characteristics of a p-n junction in reverse bias  
(b) reverse bias characteristic curve

Note that the junction current is comparatively much less in reverse bias for all voltages below the breakdown voltage. And at breakdown voltage, the current increases rapidly for a small increase in voltage. Moreover, comparison of Fig. 28.9 (b) and 28.10 (b) reveals that a p-n junction diode offers low resistance when it is forward biased and high resistance when reverse biased. At breakdown voltage in reverse biased p-n junction diode, the sharp increase in reverse current is due to sudden decrease in resistance offered by the junction.

From this we may conclude that a p-n junction diode conducts in only one direction, i.e. has unidirectional conduction of current, with electrons flowing from the n-type region to p-type end in forward bias. Some of the applications of p-n junction diode:

1. The unidirectional conducting-property of a diode is used to convert ac voltage into dc voltage as a rectifier. Diodes are also used in adaptors to recharge batteries of cell phones, CD players, laptops, etc. You will study about it in detail in the applications part.
2. A device that uses batteries, often contains a diode, as it simply blocks any current from leaving the battery, if it is reverse biased. This protects the sensitive electronics in the device.

### Intext Questions 28.5

1. Explain the concept of knee voltage.
2. (a) The knee voltage in case of silicon diode is \_\_\_\_\_ whereas in germanium diode it is \_\_\_\_\_.  
(b) In a p-n junction diode, the current flows only in \_\_\_\_\_ direction.  
(c) The reverse saturation current is of the order of \_\_\_\_\_ for germanium diodes.
3. Choose the correct option:  
(a) The I-V characteristics of a p-n junction diode in forward bias show  
(i) a non-linear curve  
(ii) linear curve  
(iii) linear as well as non-linear portions  
(iv) none of above






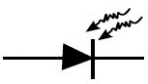
- (b) When a p-n junction is forward biased and the voltage is increased, the rapid increase in current for relatively small increase in voltage occurs
- almost immediately
  - only when the forward bias exceeds the potential barrier
  - when there is breakdown of the junction
  - none of the above

## 28.6) TYPES OF DIODES

By adjusting the levels of doping, doping material, and the geometry (size, area etc.) of a p-n junction diode, one can tune its electrical and optical behavior. Various types of diodes, their schematic symbol and their functions are given in the following Table-28.1.

**Table - 28.1 : Diodes, their schematic symbols and their functions.**

Name	Symbol	Construction	Principle mechanism	Main	Main use function
Zener diode		p-n junction diode with heavily doped p-&n-regions. Very narrow depletion layer (< 10 nm).	Zener breakdown mechanism	Provides continuous current in reverse breakdown voltage region without being damaged.	Voltage stabilization or regulation
Photo-diode		p-n junction diode. Uses light (or photo) emitting semiconductor materials, with very thin p-region, whose thickness is determined by wave length of radiation to be detected.	Photovoltaic effect into electrical current in	Converts an optical input controls in VCR & TV reverse bias.	Receivers for remote
LED		p-n junction diode with materials having band energies corresponding to near infrared region or visible light region (GaAsP or InP)	Electroluminescent	Changes an electrical input to a light output in forward bias.	Used in multimeters, digital watches, instrument displays, calculators, switch boards, burglar alarm and remote control devices

Solar cell		p-n junction diode in which either p or n region is made very thin to avoid significant absorption of light before reaching the junction.	Photovoltaic effect	Conversion of solar energy into electrical energy	<ol style="list-style-type: none"> <li>1. In satellites to power systems.</li> <li>2. To charge batteries.</li> <li>3. Calculators</li> </ol>
------------	---	---	---------------------	---	---

### Intext Questions 28.6

1. Choose the correct option
  - (a) A Zener diode is operated in
    - (i) Forward bias
    - (ii) Reverse bias
    - (iii) Both of the above
    - (iv) None of the above
  - (b) Zener diode is
    - (i) A highly doped p-n junction diode
    - (ii) A lowly doped p-n junction diode
    - (iii) A moderately doped p-n junction diode
    - (iv) Another name of normal p-n junction diode
  - (c) A Zener diode is used as a
    - (i) amplifier
    - (ii) rectifier
    - (iii) constant current device
    - (iv) constant voltage device

## 28.7 APPLICATIONS OF SEMICONDUCTOR DEVICES

On account of the miniature size and special electrical properties, diodes and triode devices find applications in almost every household appliance and gadget like gas lighter, security alarm, radio, TV, telephone, tape recorder, CD player, computer, fan regulator, emergency lights etc.

### 28.7.1 Diode (p-n junction) as a rectifier

Electricity supply in our homes provides us ac voltage and it is a sinusoidal signal of frequency 50 Hz. It means that voltage (or current) becomes zero twice in one cycle, i.e., the waveform has one positive and other negative half cycle varying symmetrically around

zero voltage level. The average voltage of such a wave is zero. Let us now learn the mechanism to convert an ac into dc.

- (a) **Half-Wave Rectification :** The signal from ac mains is fed into a step down transformer  $T$  which makes it available at the terminals  $X$  and  $Y$  (refer Fig. 28.11). The load resistance  $R_L$  is connected to these terminals through a p-n junction diode  $D$ . We use step-down transformer as most of the devices require voltage

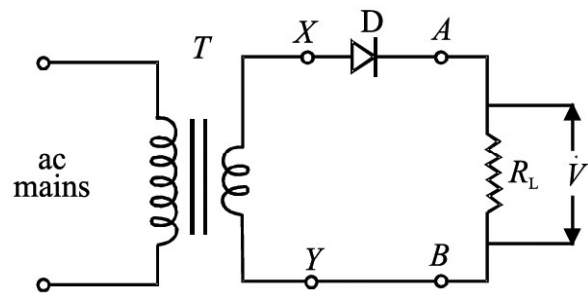


Fig. 28.11 : Half-wave rectifier circuit

levels lower than 220 V. The stepped down ac-signal is obtained at the output of step down transformer. The potential at terminal  $X$  with respect to  $Y$  will vary as a sine function with time, as shown in Fig. 28.12 (a). In the positive half cycle, during the time interval  $0$  to  $T/2$ , diode  $D$  will be forward biased and conduct, i.e., current flows

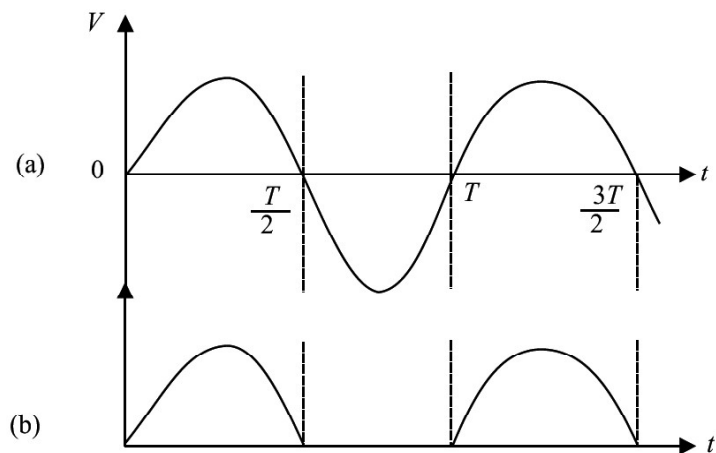


Fig. 28.12 : (a) Input ac-voltage  
(b) half wave rectified output

through  $R_L$  from  $A$  to  $B$ . However, during the negative half cycle, i.e., in the interval  $T/2$  to  $T$ ,  $D$  is reversed biased and the junction will not conduct, i.e. no current flows through  $R_L$ . This is shown in Fig. 28.12 (b). Since the p-n junction conducts only in one-half cycle of the sine wave, it acts as a half-wave rectifier. During the non-conducting half cycle, the maximum reverse voltage appearing across the diode is equal to the peak ac voltage  $V_m$ . The maximum reverse voltage that a diode can oppose without breakdown is called its **Peak Inverse Voltage (PIV)**. For rectification, it is necessary to choose a diode having PIV greater than the peak ac voltage to be rectified by it; otherwise it will get damaged. The dc voltage,  $V_{dc}$  across  $R_L$ , as measured by voltmeter in case of half-wave rectifier, is given by

$$V_{dc} = V_m/\pi \quad (28.1)$$

Where  $V_m$  is the peak ac voltage. The dc current  $I_{dc}$  through the load resistance  $R_L$  is given as:

$$I_{dc} = \frac{V_{dc}}{R_L} = \frac{V_m}{\pi R_L} \quad (28.2)$$

Note that in this case, we are utilizing only half of the input power and obviously it is not an efficient way of obtaining dc. You may logically think that instead of

one, we should use two diodes in such a way that they conduct in alternate cycles. This is known as full-wave rectification. Let us learn about it now.

- (b) **Full-Wave Rectification :** For full-wave rectification, we feed the input signal in a centre tapped step down transformer. (It has two identical secondary windings connected in series.)  $D_1$  and  $D_2$  are two p-n junction diodes, as shown in Fig. 28.13. One end of the load resistance  $R_L$  is connected to the central point Y of the secondary windings and the other end is connected to the cathode terminals of the diodes  $D_1$  and  $D_2$ . The anodes of these diodes are connected respectively to the ends X and Z of the secondary windings. The potentials at the ends X and Z are in opposite phase with respect to Y, i.e., when potential of X is positive, Z will be negative and vice versa. It is shown graphically in Fig. 28.14 (a) and (b).

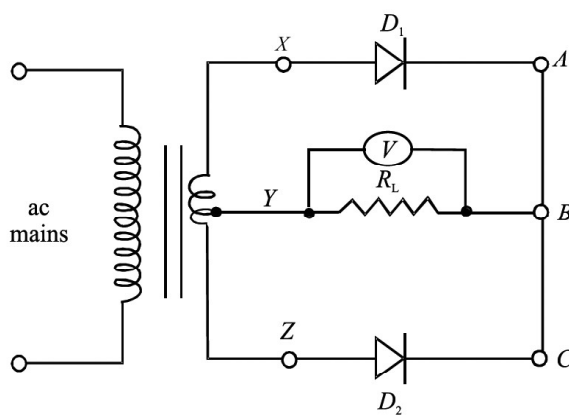


Fig. 28.13 : A full wave rectifier circuit (using two diodes)

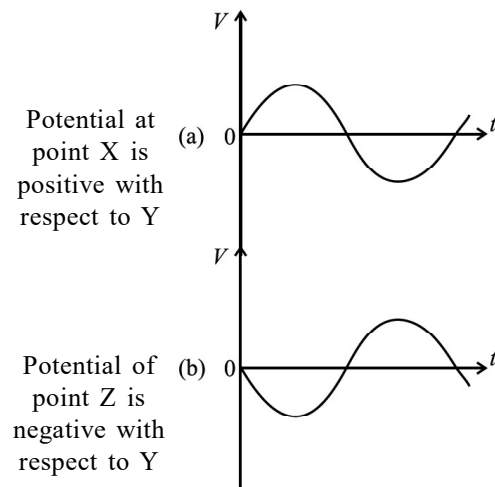


Fig. 28.14 : Potential Curves

Suppose that to start with, terminal X is positive and Z is negative with respect to Y. In this condition, diode  $D_1$  will conduct but  $D_2$  will not conduct. The current will flow through the load from B to Y and the output voltage across  $R_L$  is as shown in Fig. 28.15 (a). During the next half cycle, terminal X will be negative and Z will be positive. Under this condition, diode  $D_2$  conducts and current will again pass through the load resistance in the same direction that is from B to Y. The corresponding waveform is shown in Fig. 28.15 (b). The net output across  $R_L$  is pulsating, as shown in Fig. 28.15 (c). Since current through the load now flows over the entire cycle of the sine wave, this is called full-wave rectification. The dc voltage  $V_{dc}$  and dc current  $I_{dc}$  are given by

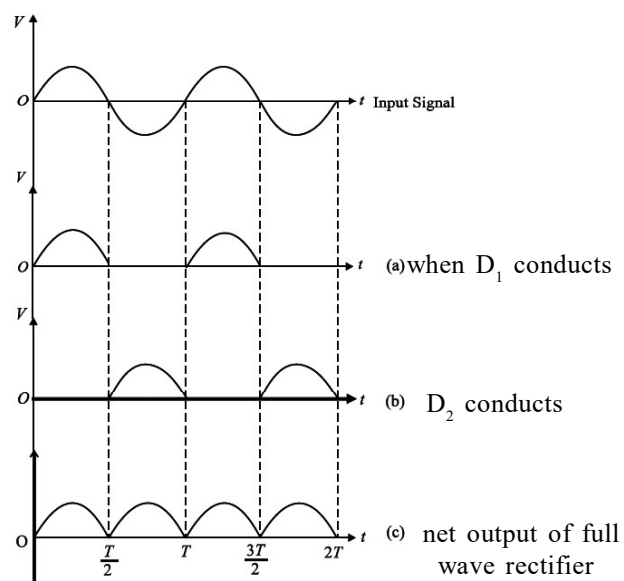


Fig. 28.15 : Voltage across  $R_L$

$$V_{dc} = 2 \times V_m / \pi \quad (28.3)$$

$$I_{dc} = \frac{V_{dc}}{R_L} = \frac{2V_m}{\pi R_L} \quad (28.4)$$

Note that the unidirectional current flowing through the load resistance after full-wave rectification pulsates from maximum to minimum (zero) and is not useful for any practical application. To reduce the fluctuating component and obtain more steady current, we filter the pulsating part. For this, a capacitor  $C$  connected across the load resistance, as shown in Fig. 28.16, filters out high frequency component.

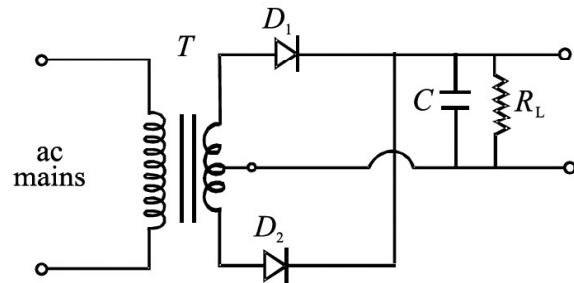


Fig. 28.16 : Circuit diagram for capacitor filter in full wave rectification

The capacitor gets charged to nearly maximum potential  $V_m$  when diode  $D_1$  conducts for period  $t = T/4$ . When the current tends to decrease for  $T/4 < t < T/2$ , the capacitor discharges itself and tries to maintain current through the load, reducing fluctuations considerably, as shown in Fig. 28.17. The larger the value of capacitor and the load resistance, the lower will be the fluctuations in the rectified dc.

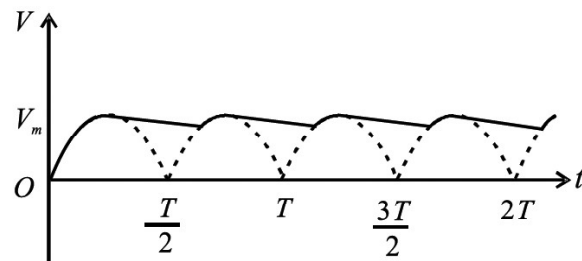


Fig. 28.17 : Output voltage when capacitor is used to filter ac

The capacitor  $C$  connected across the load to reduce fluctuations is called a *filter capacitor*. In a power supply, we use  $C$  or  $L$  or  $LC$  (or  $L$ ) and  $C-L-C$  (or  $\pi$ ) filters to reduce the rippling effect.

### Intext Questions 28.7

1. Draw the circuit of full wave rectifier with a filter capacitor.

## 28.8) SPECIAL PURPOSE OF p-n JUNCTION DIODE (ZENER DIODE)

Zener diode is a special purpose semiconductor diode, named after its inventor C. Zener in the year 1934. It is designed to operate under reverse bias in the break down region. Zener diode generally used as voltage regulator. The symbol of Zener diode is shown in the Fig. 28.18.

Two types of breakdowns observed in a Zener diode are:

**Avalanche Breakdown:** Avalanche breakdown occurs at high reverse voltage. This happens in both the normal diode as well as Zener diode. If high value of reverse voltage is applied to p-n junction, then the free electrons get sufficient energy and attain high velocity.

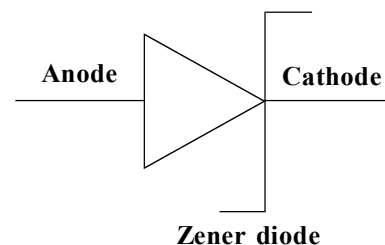


Fig. 28.18 : Zener diode symbol

These high-velocity electrons collide with the other atoms. Due to the constant collision of the electrons with new atoms, a large number of free electrons have generated and finally result into rapid increase of electric current in the diode. This kind of sudden jump in the electric current destroys the diode. Zener diode is designed to operate under this kind of breakdown voltage. If the applied reverse voltage ( $V_z$ ) is higher than 6V, the avalanche breakdown happens in the Zener diode.

**Zener Breakdown :** When the reverse bias voltage applied reaches closer to the Zener voltage, the depleting electric field in the region gets stronger enough to pull electrons from the valence band. These valence electrons gain high amount of energy from a strong electric field of the depletion region and breaks free from the parent atom. Therefore, in the Zener breakdown region, a small increase in the voltage results in a sudden increase of the electric current.

### 28.8.1 V-I Characteristics of Zener Diode

The V-I characteristics of the Zener diode is shown in the Fig. 28.19. From the graph it is evident that the Zener diode behaves like an ordinary diode when it is connected in forward bias. But when the reverse voltage is applied across the Zener diode, such that the reverse voltage rises beyond a particular value, breakdown occurs in the Zener diode.

The electric current increases sharply in the reverse direction at the breakdown voltage of the Zener diode.

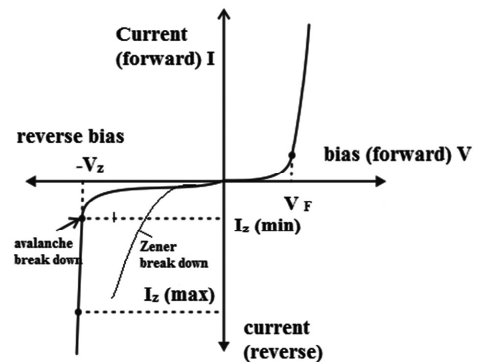


Fig. 28.19 : I-V characteristics of Zener diode.

### 28.8.2 Zener Diode as a Voltage Regulator

The half-and full-wave rectifiers with filters are the simplest type of power supplies. These provide almost pure dc- but have one deficiency. When load current is increased by decreasing resistance, the output voltage drops. This is because, when large current is drawn, the filter capacitor gets discharged more and its voltage across the load resistor reduces. Similarly, if the ac- input changes, the dc-output voltage also varies. Obviously, a supply with varying output voltage affects the performance of different devices being operated with it. For example, if we operate an amplifier, the quality of sound reproduced by it will get deteriorated. To remove this deficiency, a Zener diode is used with simple power supplies which give constant dc voltage. Such a circuit is called regulated power supply. The Zener regulated voltage supply circuit is shown in Fig. 28.20. It consists of a Zener diode with breakdown voltage  $V_z$ . This will be equal to the stabilized output voltage  $V_o$ . A suitable series resistance  $R_s$  is included to control circuit current and dissipate excess voltage. The anode of Zener diode is connected to the negative terminal of input

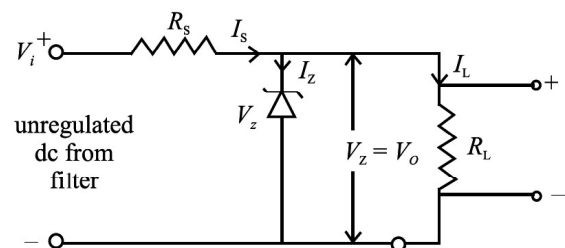


Fig. 28.20 : Zener diode as a stabilizer

supply, and the cathode is connected in series with  $R_s$  to positive terminal of input supply, that is, the Zener is connected in reverse bias condition. The load resistance is connected across the Zener diode. The Zener regulator will only operate if the input supply voltage to the regulator,  $V_i$  is greater than  $V_z$ . After breakdown, the voltage across it remains nearly constant and is independent of the current passing through it. The current  $I_s$  passing through  $R_s$  is given by the equation

$$I_s = (V_i - V_z)/R_s \quad (28.5)$$

This current divides in two parts: the Zener current  $I_z$  and load current  $I_L$ . Applying Kirchoff's current law, we can write

$$I_s = I_z + I_L$$

or 
$$I_z = I_s - I_L \quad (28.6)$$

For Zener diode to operate, some current  $I_{z\min}$  should always flow through it. Therefore, the load current  $I_L$  should always be less than the main current  $I_s$ . Typical value of minimum  $I_z$  may range from 5 mA to 20 mA.

If load current is zero, the entire  $I_s$  will pass through Zener diode and output voltage  $V_o$  will be equal to  $V_z$ . When some load current is drawn, say  $I_L$ , the Zener current will decrease by the same amount but the output voltage will remain  $V_z$ . Similarly, if the ac main voltage increases or decreases, the input voltage,  $V_i$  will increase or decrease accordingly. It will result in change of  $I_s$  given by Eqn. (28.5). Due to change in  $I_s$ , the change in  $V_i$  will appear as a drop across the series resistance  $R_s$ . The Zener voltage  $V_z$  and hence  $V_o$  will remain unchanged. Thus we see that the output voltage has been stabilized against the variations in the current and the input voltage. The power dissipation in Zener diode is given by the relation:

$$P_d = V_z \times I_z \quad (28.7)$$

This dissipation should not exceed the maximum power dissipation rating recommended by the manufacturer for Zener diode. Let us now understand the design of a Zener regulated power supply with the following example.

### Example 28.1

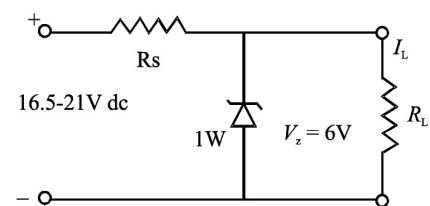
The load current varies from 0 to 100 mA and input supply voltage varies from 16.5 V to 21 V in a circuit. Design a circuit for stabilized dc supply of 6 V.

#### Solution :

We choose a Zener diode of 6 V. Let  $I_{z\min}$  be 5 mA. The maximum current will flow through the Zener when there is no load current. Its magnitude will be  $(100 + 5) \text{ mA} = 0.105 \text{ A}$ .

The value of  $R_s$  is determined by the minimum input voltage and maximum required current

$$R_s = \frac{V_{z\min} - V_z}{I_{\max}} = \frac{16.5 \text{ V} - 6 \text{ V}}{105 \text{ mA}} = 100 \text{ ohms}$$



The current through the Zener diode will be maximum when the input voltage is maximum, that is 21 V and  $I_L = 0$ .

Therefore, the maximum Zener current  $I_{\max} = (21V - 6V)/100 \Omega = 0.15 A$ .

The maximum power dissipation in the diode is  $6 V \times 0.15A = 0.9 W$ .

It means that we should use a Zener diode of 6 V, 1 W and resistance  $R_s$  of 100  $\Omega$ . It should be connected in the circuit as shown above. It will give a stable output of 6 V for the specified ranges of load and input variation.

### 28.8.3 Opto-electronic junction devices

Optoelectronic junction devices are special types of semiconductor devices. These devices are able to convert light energy to electrical energy or vice versa. Solid crystalline materials, which are heavier than insulators but lighter than metals, are used to make optoelectronic devices. Some of the optoelectronic devices are:

- a. Photodiodes (used for detecting optical signal)
- b. Light emitting diodes (LED) (which converts electrical energy into light)
- c. Photovoltaic devices (which converts optical radiation into electricity, also called solar cells)

(a) **Photodiodes** : It is a device which converts light energy into electric energy. It works under reverse bias conditions. Materials such as Silicon, Germanium, and Indium gallium arsenide are used to prepare the photodiodes. Photodiodes can be represented using the following symbol [Fig. 28.21 (a)].



Fig. 28.21 (a) : Photodiode symbol

Here, it should be noted that the symbol for the photodiode is the same as LED. The arrows point inwards means the photodiode absorbs light energy whereas arrows outwards means LED emits light energy.

#### Working of photodiodes :

A Photodiode is a special p-n junction diode, fabricated with a transparent window to allow light ( $h\nu$ ) to fall on the diode [see Fig. 28.21 (b)]. The working procedure of the photodiode is given point wise for easy understanding of this concept.

- i. It is operated under reverse bias.
- ii. Photodiode is illuminated with photon energy ( $h\nu$ ), greater than the energy gap ( $E_g$ ). Then, electron-hole pairs are generated due to the absorption of photons, near the depletion region of the diode, as shown in the Fig. 28.21 (b).
- iii. On account of applied electric field, electrons and holes are separated before they recombine.

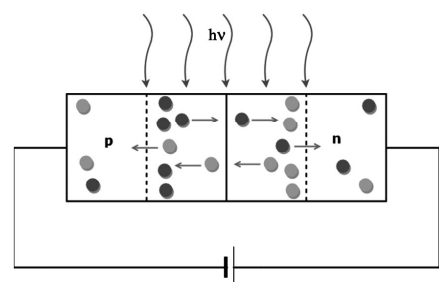


Fig. 28.21 (b) : Illuminated photodiode under reverse bias



- iv. The direction of the electric field makes the electrons to reach n-side and holes to reach p-side.
- v. Due to the electric field, electrons are collected on n-side and holes are collected on p-side giving rise to an emf.
- vi. If an external load is connected, then current flows. The magnitude of the photocurrent depends on the intensity of incident light.
- vii. Photocurrent is proportional to incident light intensity.

Based on the above points it is evident that a change in the current causes change in the light intensity under reverse bias condition. A typical I-V characteristic of photodiode is shown in the Fig. 28.21 (c). This principle allows the photodiode to work as a photodetector to detect optical signals.

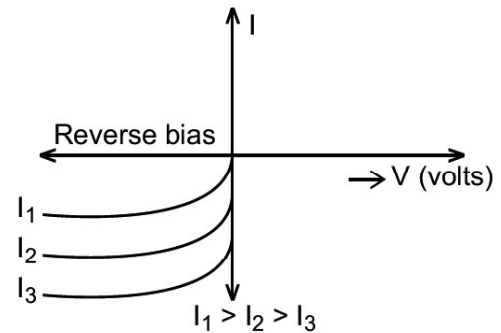


Fig. 28.21 (c) : I-V Characteristics of photodiode under reverse bias, for different light (illuminated) intensities

- (b) **Light emitting diodes** : A Light Emitting Diode (LED) is also a special type of p-n junction diode. The light emitting diode is specially doped semiconductor and emits light in the forward biased state. Aluminum indium gallium phosphide (AlInGaP) and indium gallium nitride (InGaN) are two of the most commonly used semiconductors for LED technologies. A schematic diagram of LED is shown in the Fig. 28.22 (a).

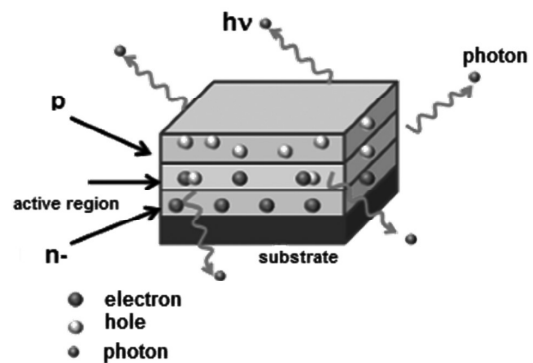


Fig. 28.22 (a) : Schematic diagram of LED

Some of the important features of working conditions of LED are given below:

- i. LED is a heavily doped p-n junction.
- ii. Under forward bias condition it emits spontaneous radiation.
- iii. It is encapsulated with a transparent cover so that emitted light can come out.
- iv. If LED is forward biased, then electrons are sent from  $n \rightarrow p$  and holes from  $p \rightarrow n$ .
- v. At the junction boundary the concentration of minority carriers increases compared to the equilibrium concentration (or no bias condition).
- vi. On recombination, the energy is released in the form of photons.
- vii. Photons with energy equal to or slightly less than the band gap are emitted.
- viii. If the forward current of the diode is small, then the intensity of light emitted is small.

- ix. As the forward current increases, intensity of light increases and reaches a maximum. With further increase in the forward current, it results in decrease of light intensity.

The I-V characteristics of a LED is similar to the Si-junction diode. But, the threshold voltages are much higher, compared to junction diodes. Threshold voltages are slightly different for each color. The reverse breakdown voltages of LEDs are very low, typically around 5V.

Red, yellow, orange, green and blue LEDs are known as commercially available LEDs. The visible LEDs must at least have a band gap of 1.8 eV. Gallium Arsenide – Phosphide ( $\text{GaAs}_{1-x}\text{P}_x$ ) is used for colour LEDs.  $\text{GaAs}_{0.6}\text{P}_{0.4}$  ( $E_g \sim 1.9$  eV) is used for red LED, GaAs ( $E_g \sim 1.4$  eV) is used for making infrared LED. LEDs are used in remote controls, burglar alarm systems and optical communication systems.

LEDs have many advantages, when compared to the incandescent low power lamps. They have low operational voltage and less power. The bandwidth of emitted light is in the range of 100 Å to 500 Å or in other words they are monochromatic. They have long life and ruggedness. In addition they have fast on-off switching capability.

- (c) **Solar cells or photovoltaic (PV) cell** : A solar cell is basically a p-n junction, which generates emf when solar radiation falls on the p-n junction. It works on the principle of photovoltaic effect. If a diode is illuminated by photons (light) with energy ( $h\nu > E_g$ ), then electron-hole pairs are generated due to the absorption of photons. For more power, one should maintain larger junction area and no external bias condition.

A solar cell is defined as a device whose electrical characteristics such as current, voltage, or resistance should vary when exposed to light. Fig. 28.22 (b) shows typical p-n junction solar cell.

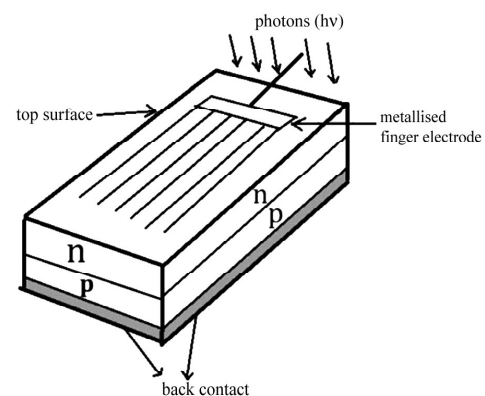


Fig. 28.22 (b) : Solar cell (p-n junction)

Silicon wafer of p-type about 300  $\mu\text{m}$  is taken over which a thin layer ( $\sim 0.3$   $\mu\text{m}$ ) of n-type silicon is grown on one-side by diffusion process. The other side of p-type silicon (Si) is coated with a metal (back contact). On the top of n-Si layer, metal finger electrode (or metallic grid) is deposited. This acts as a front contact. The metallic grid occupies only a very small fraction of the cell area ( $< 15\%$ ) so that light can be incident on the cell from the top.

The generation of emf by a solar cell happens when light falls on. It is due to the following three basic processes: (i) generation (ii) separation and (iii) collection. Generation of e-h pairs due to light (with  $h\nu > E_g$ ) close to the junction; separation of electrons and holes due to electric field of the depletion region. Electrons are swept to n-side and holes to p-side; the electrons reaching the n-side are collected by the front contact and holes reaching p-side are collected by the back contact. Thus p-side becomes positive and n-side becomes negative giving rise to photovoltage. If an external load is connected, as shown

in the Fig. 28.22 (c), photocurrent ( $I_L$ ) flows through the load ( $R_L$ ). I-V characteristic of solar cell is shown in the Fig. 28.22 (d). Since solar cell does not draw current but supplies the current to the load, therefore, I-V characteristics is shown in the fourth quadrant of the graph.

Solar cell materials: Semiconductors with band gap close to 1.5 eV are known as ideal materials for solar cell fabrication. The semiconducting materials like Si ( $E_g = 1.1$  eV), GaAs ( $E_g = 1.43$  eV), CdTe ( $E_g = 1.45$  eV), CuInSe<sub>2</sub> ( $E_g = 1.04$  eV), etc. are used in the fabrication. The important criteria for the selection of a material for solar cell fabrication are as follows:

- (i) band gap ( $\sim 1.0$  to  $1.8$  eV)
- (ii) high optical absorption ( $\sim 10^4$  cm<sup>-1</sup>)
- (iii) electrical conductivity of raw materials

Sunlight is not necessarily required for a solar cell. Light with photon energies greater than the band gap will work for solar cell.

**Uses :** They are used to power electronic devices in satellites, space vehicles and power suppliers to some electronic instruments.

### Intext Questions 28.8

1. What will be the output voltage, if you connect a Zener diode in forward bias instead of reverse bias in the regulator circuit of Example 28.1?
2. Fill in the blanks
  - a) The Zener diode is based on the \_\_\_\_\_ breakdown mechanism.
  - b) A photodiode is operated in \_\_\_\_\_ bias.
  - c) In a photodiode, the p-n junction is made from \_\_\_\_\_ semiconductor material.
  - d) LED's are made up of the conductor material from \_\_\_\_\_ of the periodic table.
  - e) The light emitting diodes operate in \_\_\_\_\_ bias.
  - f) The arrow in the symbol of LED symbolizes \_\_\_\_\_ of light.
  - g) In an LED light is emitted due to \_\_\_\_\_ of electrons and holes.

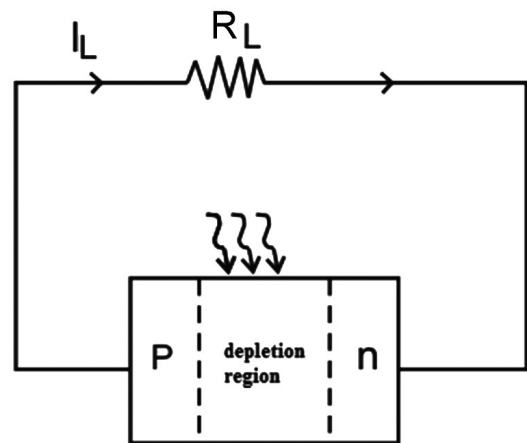


Fig. 28.22 (c) : Illuminated p-n junction solar cell

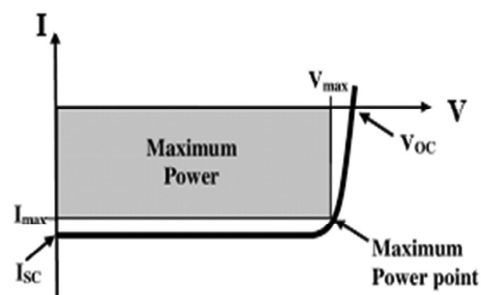


Fig. 28.22 (d) : I-V characteristics of solar cell

- h) LED is based on the principle of \_\_\_\_\_.
- i) Solar cells are based on \_\_\_\_\_ effect.
- j) When sunlight having energy \_\_\_\_\_ than the band gap energy falls on the solar cell, it is \_\_\_\_\_ and frees electron-hole pairs.

## 28.9) TRANSISTORS

In the preceding sections, you have learnt about a p-n junction diode, which permits current to flow only in one direction. This limits its applications to rectification and detection. A more useful semiconductor device is a bipolar junction transistor. The invention of transistor by John Bardeen, Walter Brattain and William Shockley in 1948 at Bell laboratory in USA revolutionized the electronic industry. The transistors find many applications in our daily life ranging from gas lighter to toys to amplifiers, radio sets and television. In the form of switching device, these can be used to regulate vehicular traffic on the roads. They form key elements in computers, space vehicles, power systems in satellites and communication.

A transistor is basically a silicon or germanium crystal containing three alternate regions of p and n-type semiconductors as shown in Fig. 28.23. These three regions are called emitter (E), base (B) and collector (C). The middle region is the base and the outer two regions are emitter and collector. Note that the emitter and collector are of the same type (p or n) and collector is the largest of the three regions.

The base terminal controls the current flowing between the emitter and the collector. This control action gives the transistor an added advantage over the diode, which has no possibility of controlling the current flow. Depending on the type of doping, the transistors are classified as n-p-n or p-n-p. In general, the level of doping decreases from emitter to collector to base.

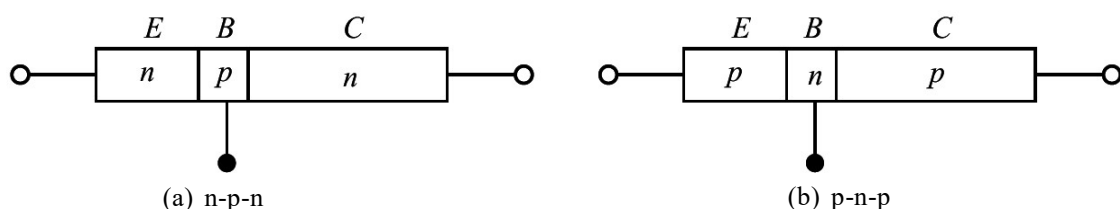


Fig. 28.23 : Types of transistors

The names of the terminals of a transistor give clear indication of their functions. In case of a n-p-n transistor, the majority carriers (electrons) from the emitter are injected into base region. Since base is a very lightly doped thin layer, it allows most of the electrons injected by the emitter to pass into the collector. Being the largest of three regions, the collector dissipates more heat compared to the other two regions.

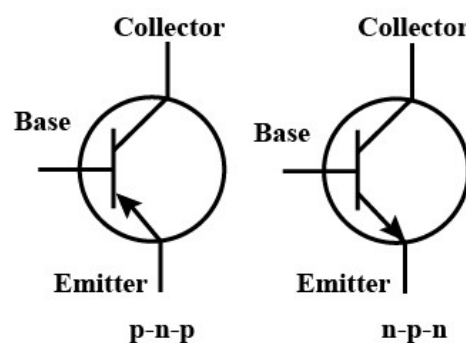
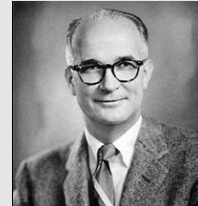


Fig. 28.24 : Symbols of p-n-p and n-p-n transistors

The symbolic representations of n-p-n and p-n-p transistors are shown in Fig. 28.24. The arrow head indicates the direction of flow of conventional current. In a n-p-n, transistor, the emitter current is due to flow of electrons from emitter to base, and the conventional current flows from base to emitter and hence the arrow head points out from the base. In case of p-n-p transistor, the emitter current comprises flow of holes from emitter to base. Thus the conventional current flows from emitter to base. Since transistors are bipolar devices, their operation depends on both the majority and minority carriers (electrons and holes).

### William Bradford Shockley (1910 – 1989)

England born, American physicist W.B. Shockley was one of the three scientists who received 1956 Nobel Prize in physics for the discovery of transistor. Basically a solid state physicist, shockley contributed significantly to the development of theoretical understanding of bands in semiconductors, order and disorder in alloys; theory of vacuum tubes, theory of dislocations and theory of ferromagnetic domains. He is truly one of the pioneers of electronic revolution.



#### 28.9.1 Working Principle

Consider an n-p-n transistor, because it is more commonly used. When no voltage is applied across the transistor, diffusion of free electrons across the junctions produces two depletion layers, as shown in Fig. 28.25. For each depletion layer, the barrier potential is about 0.7 V at 25°C for a silicon transistor and 0.3 V for a germanium transistor. As you may be aware, silicon transistors are more widely used than germanium transistors because of higher voltage rating, greater current ratings, and low temperature sensitivity. For our discussion, we refer to silicon transistors, unless otherwise indicated.

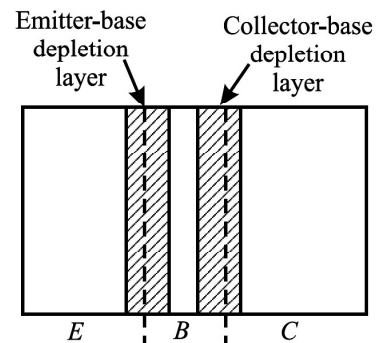


Fig. 28.25 : Depletion layers in a transistor when no voltage is applied

Since the three regions in a transistor have different doping levels, the depletion layers have different widths. If a region is heavily doped, the concentration of ions near the junction will be more, resulting in thin depletion-layer and vice versa. Since the base is lightly doped as compared to emitter and collector, the depletion layers extend well into it, whereas the penetration in emitter/collector regions is to a lesser extent.

#### I. n-p-n Transistor

A typical biasing scheme of a n-p-n transistor is shown in Fig. 28.26 (a). Note that *the emitter-base junction is forward biased while the collector-base junction is reversing biased*. We therefore expect a large emitter current and low collector current. But in practice, we observe that the collector current is almost as large as the emitter current. Let us understand the reason. When forward bias is applied to the emitter junction, free electrons in the emitter have to overcome the barrier potential to enter the base region [see Fig. 28.26 (b)]. When  $V_{BE}$  exceeds barrier potential (0.6 to 0.7V for silicon transistor), these electrons enter the base region, as shown in Fig. 28.26 (c). Once inside the base, these electrons can flow

either through the thin base into the external base lead or across the collector junction into the collector region. The downward component of base current is called *recombination current*. It is small because the base is lightly doped and only a few holes are available. Since the base region is very thin and it receives a large number of electrons, for  $V_{BE} > 0.7V$ , most of these electrons diffuse into the collector depletion layer. The free electrons in this layer are pushed (by the depletion layer field) into the collector region [(Fig.28.26 (d)] and flow into the external collector lead. So, we can say that a steady stream of electrons leaves the negative source terminal and enters the emitter region. The forward bias forces these electrons to enter the base region. Almost all these electrons diffuse into the collector depletion layer through the base. The depletion layer field then pushes a steady stream of electrons into the collector region. In most transistors, more than 95 percent emitter-injected electrons flow to the collector; less than 5 percent flow to the external base.

The key to transistor actions is due to the lightly doped thin base sandwiched between the heavily doped emitter and the intermediately doped collector. Free electrons passing through the base stay in the base for a short time and reach the collector. The relation between collector current ( $I_C$ ) and emitter current ( $I_E$ ) is expressed in terms of large signal current gain,  $\alpha$ , of a transistor. It is defined as:

$$\alpha = \frac{I_C}{I_E} \quad (28.8)$$

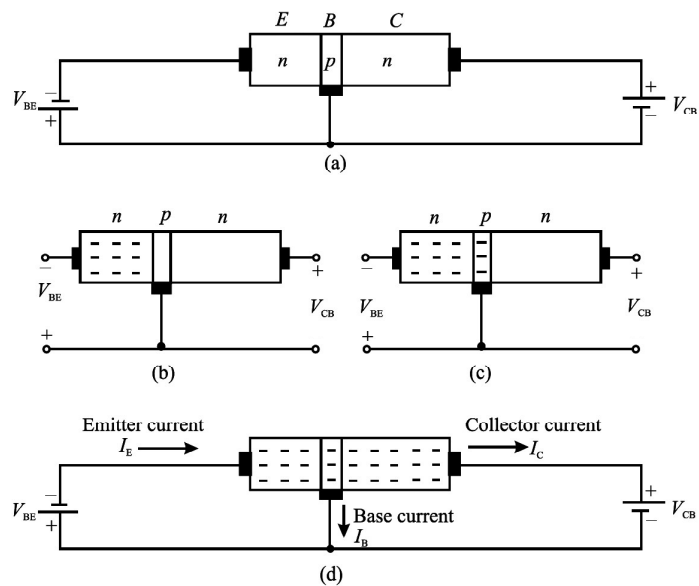
Here,  $\alpha$  is nearly equal to one, but always slightly less than one.

Similarly, one can relate the collector current to the base current in a transistor. It is denoted by Greek letter beta:

$$\beta = \frac{I_C}{I_B} \quad (28.9)$$

Beta signifies the current gain of the transistor in common-emitter configuration. The value of  $\beta$  is significantly greater than one. Since emitter current equal to the sum of collector current and base current, we can write

$$I_E = I_C + I_B$$



**Fig. 28.26 :** A n-p-n transistor when (a) emitter is forward-biased and collector is reverse-biased, (b) free electrons in an emitter, (c) free electrons injected into base; and (d) free electrons pass through the base to the collector.

On dividing throughout by  $I_C$ , we get

$$\frac{I_E}{I_C} = 1 + \frac{I_B}{I_C} \quad (28.10)$$

In terms of  $\alpha$  and  $\beta$ , we can rewrite it as

$$\frac{1}{\alpha} = 1 + \frac{1}{\beta}$$

$$\beta = \frac{\alpha}{1-\alpha} \quad (28.11)$$

## II. A p-n-p Transistor

In this section one can understand how a p-n-p transistor differs from a n-p-n transistor in detail. A p-n-p transistor biased for operation in the active region is shown in Fig. 28.27. Note that we reverse the battery terminals when n-p-n transistor is substituted by p-n-p transistor.

As before, the emitter-base junction is forward biased by battery of voltage  $V_{EB}$  and the collector base junction is reverse biased by a battery of voltage  $V_{CB}$ . The resistance of the emitter-base junction is very small due to its forward bias as compared to the collector-base junction (which is reverse biased). Therefore, we apply small forward bias voltage (0.7 V) to the emitter-base junction, whereas the reverse bias voltage applied to the collector-base junction is of much higher value (9 V). The forward bias of emitter-base junction makes the majority carriers that are the holes, in emitter (p-region), to diffuse to the base (n-region), on being repelled by the positive terminal of the battery. As width of the base is extremely thin and it is lightly doped, very few (two to five percent) of total holes that enter the base recombine with electrons and 95% to 98% reach the collector region. Due to reverse bias of the collector-base region, the holes reaching this region are attracted by the negative potential applied to the collector, thereby increasing the collector current ( $I_C$ ). Therefore, increase in emitter current ( $I_E$ ) increases collector current and equations (28.8) – (28.11) holds in this case as well.

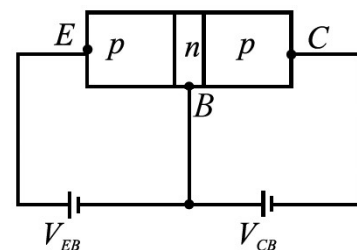


Fig. 28.27 : A p-n-p transistor biased for active operation

### Intext Questions 28.9

1. Choose the correct option :
  - a) The arrow head in the symbol of a transistor points in the direction of
    - (i) hole flow in the emitter region
    - (ii) electron flow in emitter region

- (iii) majority carriers flow in the above region  
 (iv) none of the above
- b) The emitter current in a transistor in normal bias is
- (i) less than the collector current  
 (ii) equal to sum of base current and collector current  
 (iii) equal to base current  
 (iv) none of the above
2. Fill in the blanks
- (a) A transistor has \_\_\_\_\_ regions and \_\_\_\_\_ junctions.  
 (b) In a transistor, \_\_\_\_\_ has the least thickness.  
 (c) The emitter region is \_\_\_\_\_ doped, whereas \_\_\_\_\_ region has the least doping.  
 (d) The collector of the transistor has \_\_\_\_\_ size and \_\_\_\_\_ doping.  
 (e) The transistor is said to be in active region when \_\_\_\_\_ junction is forward biased and \_\_\_\_\_ junction is reverse biased.  
 (f) The two types of transistors are \_\_\_\_\_ and \_\_\_\_\_.

### 28.9.2 Transistor Configurations

A transistor is a two-port device; it can take an input and deliver an output. For both input and output, two terminals are needed. This can be done in a transistor by making one of the three terminals as common. The configurations of a transistor in which one of the terminals is common to both input and output are shown in Fig. 28.28.

- When emitter is common to both input and output circuits, we obtain common emitter (CE) configuration [Fig. 28.28 (a)];
- When base is common to both input and output circuits, we obtain common base (CB) configuration [Fig. 28.28 (b)]; and
- When collector is common to both input and output circuits, we have common collector (CC) configuration [Fig. 28.28 (c)].

In each of these configurations, the transistor characteristics are unique. The CE configuration is used most widely because it provides voltage, current and power gains. In the CB configuration, the transistor can be used as a constant current source while the CC configuration is usually used for impedance matching

For each configuration, we can plot three different characteristics:

- Input characteristics,
- Output characteristics, and
- Transfer characteristics, depending on the nature of quantities involved.



Table-28.2 gives various quantities related to each of these characteristics in all the three configurations and the transistor constants of interest.

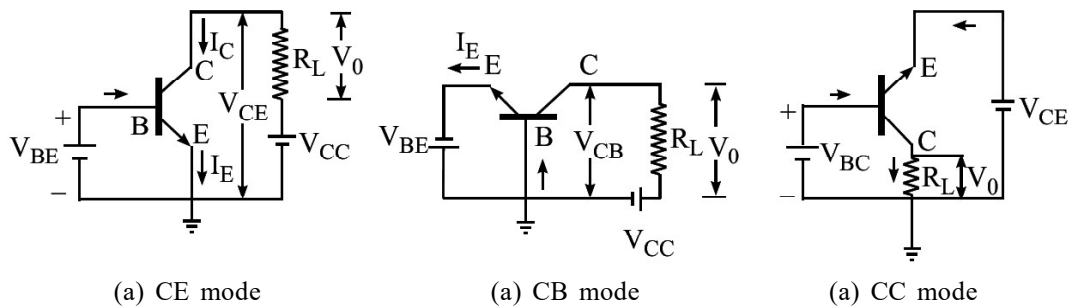


Fig. 28.28 : Transistor configuration

Table-28.2: Physical quantities of interest in different characteristics of a transistor

Configuration	Input Characteristic	Output Characteristic	Transfer Characteristic	Important transistor constant
CE	$V_{BE}$ and $I_B$ with $V_{CE}$ as parameter	$V_{CE}$ and $I_C$ with $I_B$ as parameter	$I_B$ and $I_C$	Current gain, $\beta$
CB	$V_{BE}$ and $I_E$ with $V_{CB}$ as parameter	$V_{CB}$ and $I_C$ with $I_E$ as parameter	$I_E$ and $I_C$	Current gain, $\alpha$
CC	$V_{CB}$ and $I_B$ with $V_{CE}$ as parameter	$V_{CE}$ and $I_E$ with $I_B$ as parameter	$I_B$ and $I_E$	

To work with a transistor, you will be required to identify its base, emitter and collector leads. To do so, you can follow the following steps.

- (i) Look for a small notch provided on the metallic cap. The terminal close to the notch is emitter. To identify other two terminals, turn the transistor up-side-down.
- (ii) Identify the base and the collector as shown in Fig. 28.29.
- (iii) Like a p-n junction diode, transistors are also designated with two letters followed by a serial number.

The first letter gives an indication of the material. A is for germanium and B is for silicon. The second letter indicates the main application: C is used for audio frequency transistors, D for power transistors and F for radio-frequency transistors. The serial number consists of digits assigned by the manufacturer for identification. For example, AC 125 represents germanium transistor for audio frequency (AF) applications.

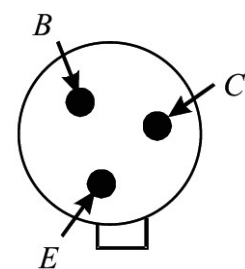


Fig. 28.29 : Identifying transistor leads

### 28.10 TRANSISTOR CHARACTERISTICS

As mentioned earlier, operation of a transistor can be studied with input and output I-V characteristics. The nature of these characteristics is unique and depends on the configuration used. Let us first study CE configuration.

### 28.10.1 Common Emitter (CE) Configuration of a n-p-n Transistor

Common emitter characteristics of a transistor relate voltage and current when emitter is common to both input and output circuits. The circuit diagram for CE characteristics of a n-p-n transistor is shown in Fig. 28.30.  $V_{BB}$  is a variable dc supply of 0-3 V and  $V_{CC}$  is a variable dc supply of 0-15 V.  $R_B$  is a base resistor. It is used to control base to emitter voltage,  $V_{BE}$ .  $R_L$  is a load register connected in series in the collector circuit.

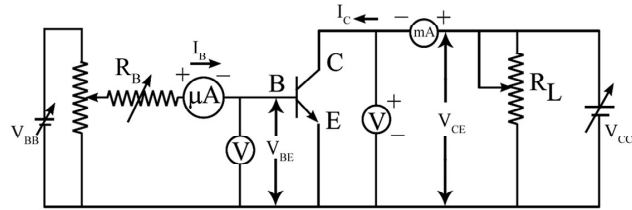


Fig. 28.30 : Circuit diagram for input and output characteristics of a transistor in CE configuration.

#### Input characteristics

In CE configuration, the input characteristics show the variation of  $I_B$  with  $V_{BE}$  when  $V_{CE}$  is held constant. To draw this characteristics,  $V_{CE}$  is kept at a suitable value using collector supply  $V_{CC}$ . Then  $V_{BE}$  is changed in steps and corresponding values of  $I_B$  are measured with the help of micro-ammeter, connected to base. Fig. 28.31 shows typical input characteristics of a n-p-n transistor in CE configuration.

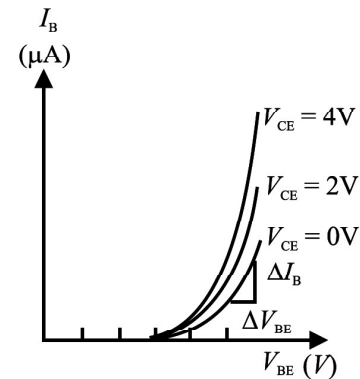


Fig. 28.31 : Input characteristics of a typical npn- transistor in CE configuration

Note that for a given value of  $V_{CE}$ , the curve is as obtained for forward biased p-n junction diode. For  $V_{BE} < 0.5$  V, there is no measurable base current ( $I_B = 0$ ). However,  $I_B$  rises steeply for  $V_{BE} > 0.6$ V.

From the reciprocal of the slope of input characteristic, we get input resistance of the transistor defined as the ratio of small change in base - emitter voltage to the small change produced in the base current at constant collector - emitter voltage. Usually, the value of  $R_{ie}$  is in the range 20-100  $\Omega$ . You should note that since the curve is not linear, the value of input resistance varies with the point of measurement. As  $V_{CE}$  increases, the curve tends to become more vertical and the value of

$$R_{ie} = \left. \frac{\Delta V_{BE}}{\Delta I_B} \right|_{V_{CE}} \quad (28.12)$$

#### Output characteristics

The output characteristic curves depict the variation of collector current  $I_C$  with  $V_{CE}$ , when base current  $I_B$  is kept constant. To draw output characteristics,  $I_B$  is fixed, say at 10  $\mu$ A, by adjusting  $V_{BB}$  supply.  $V_{CE}$  is then increased from 0 to 10 V in steps of 0.5 V by varying  $V_{CC}$  supply and the corresponding value of  $I_C$  is noted. Similarly, the output characteristics can be obtained at  $I_B = 20 \mu$ A, 40  $\mu$ A, 60  $\mu$ A, 80  $\mu$ A. However, in no case, the maximum base current rating

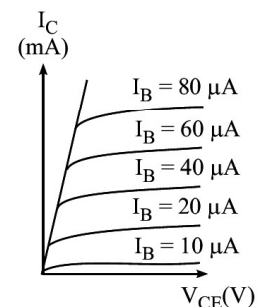


Fig. 28.32 : Output characteristics for npn translator CE configuration

of the transistor should be exceeded. The output characteristics of this configuration are shown in Fig. 28.32. From the output characteristics, you will note that  $I_C$  changes with increase in  $V_{CE}$  for a given value of  $I_B$  and  $I_C$  increases with  $I_B$  for a given  $V_{CE}$ . From these characteristics, we can calculate output admittance ( $h_{oe}$ ):

$$h_{oe} = \frac{\Delta I_C}{\Delta V_{CE}} \tag{28.13}$$

### 28.10.2 Common Emitter (CE) Configuration of a p-n-p Transistor

In the preceding section, you learnt to draw input and output characteristics of a n-p-n transistor in common emitter configuration. Now we will consider a p-n-p transistor. Fig. 28.33 shows the circuit diagram for CE characteristics of a p-n-p transistor. The transistor is biased to operate in the active region. The micro-ammeter and voltmeter are used in the base-emitter circuit to measure the base current ( $I_B$ ) and the voltage between base and emitter. Similarly, milliammeter and voltmeter are connected in collector-emitter circuit to measure the collector current ( $I_C$ ) and voltage between collector and emitter ( $V_{CE}$ ).

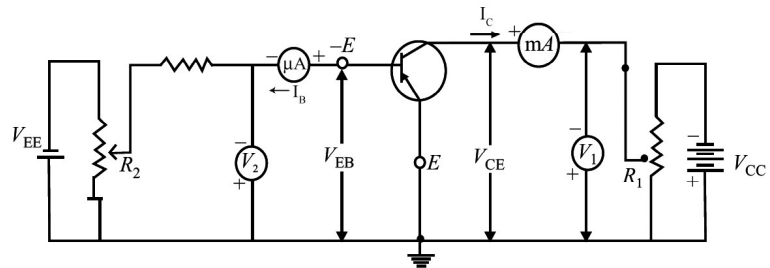


Fig. 28.33 : Circuit diagram for obtaining input and output characteristics of a p-n-p transistor in CE configuration

### Input Characteristics

Input characteristics are plotted between  $V_{BE}$  and  $I_B$  at different constant values of  $V_{CE}$ . To plot input characteristics, the potentiometer  $R_1$  in the emitter-collector circuit is adjusted till the voltmeter shows constant value. Then potentiometer in the emitter-base circuit is adjusted in such a way that base-emitter voltage is zero. For this value, base current is also observed to be zero. Keeping the  $V_{CE}$  constant,  $V_{BE}$  is increased gradually and change in base current is noted with the help of micro-ammeter. To plot input characteristics at  $V_{CE} = -2V$ , say, the potentiometer in emitter-collector circuit is adjusted till the voltmeter in the same circuit reads 2V. Then potentiometer in the emitter-base circuit is adjusted to make  $V_{BE}$  zero.

Then  $V_{BE}$  is increased gradually, keeping  $V_{CE}$  constant. Similarly the input characteristics of the transistor in the CE configuration can be drawn for different values of  $V_{CE} = -6V, -10V$  and so on.

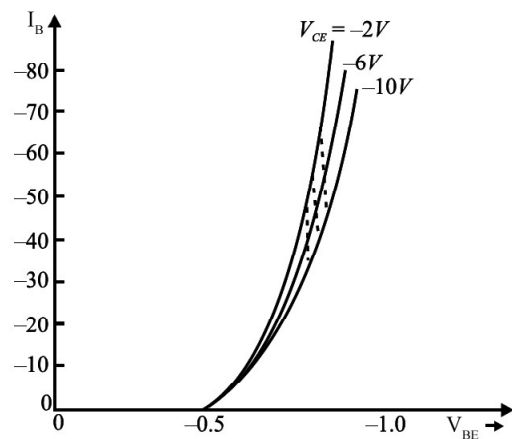


Fig. 28.34 : Input characteristics of npn transistor in CE-configuration

Fig. 28.34 shows typical input characteristics of CE-configuration. As may be noted,

the nature of *input characteristics is similar to the forward characteristics of p-n junction diode*. The base current remains zero as long as the base voltage is less than the barrier voltage (for silicon transistor, it is  $\sim 0.7$  V). As the base voltage exceeds barrier voltage, current begins to increase slowly and then rose abruptly. From the reciprocal of the slope of the curve of input characteristic, the a.c input resistance of the transistor can be calculated.

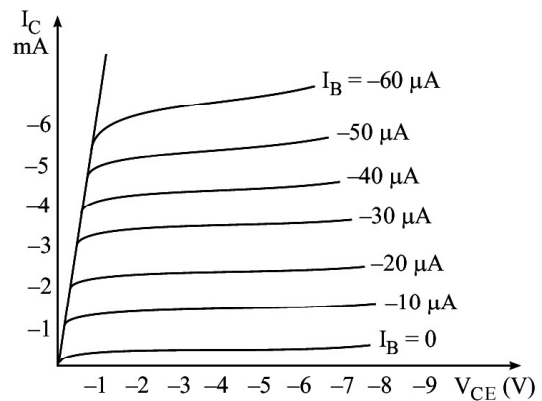
**Input resistance ( $R_{in}$ )** of the transistor in CE configuration is expressed as

$$R_{in} = \left. \frac{\Delta V_{EB}}{\Delta I_B} \right|_{V_{CE} = \text{constant}} \quad (28.14)$$

In this configuration  $R_{in}$  is typically of the order of one  $k\Omega$ .

### Output characteristics

These are the graphs between collector-emitter voltage ( $V_{CE}$ ) and the collector current ( $I_C$ ) at different constant values of base current ( $I_B$ ). To draw these characteristics,  $V_{CE}$  is made zero and  $V_{BE}$  is adjusted till the microammeter in the base-emitter circuit is set to read a constant value. Thus  $V_{BE}$  is adjusted to make  $I_B$  constant at a particular value. Now keeping  $I_B$  constant,  $V_{CE}$  is increased from zero in a number of steps and the corresponding collector current  $I_C$  is noted with the help of milliammeter connected in series with collector. To plot the output characteristics at  $I_B = 50 \mu A$ ,  $V_{BE}$  is adjusted till the microammeter reads  $50 \mu A$ . Now increase  $V_{CE}$  gradually and note corresponding values of  $I_C$ . The graph between  $V_{CE}$  and  $I_C$  gives the output characteristics at  $I_B = 50 \mu A$ . Similarly, the output characteristics can be obtained at  $I_B = 10 \mu A, 20 \mu A$  and so on.



**Fig. 28.35 :** Output characteristics of a typical pnp-transistor in CE configuration

Fig. 28.35 shows output characteristics of p-n-p transistor for CE configuration.

### Intext Questions 28.10

1. Fill in the blanks
  - (a) The \_\_\_\_\_ curve relates the input current with input voltage, for a given output voltage.
  - (b) The \_\_\_\_\_ curve relates the output current with the output voltage for a given input current.
  - (c) In common emitter configuration of a transistor, the \_\_\_\_\_ and \_\_\_\_\_ are the output terminals
  - (d) The \_\_\_\_\_ and \_\_\_\_\_ are the input terminals, whereas \_\_\_\_\_ and \_\_\_\_\_ are the output terminals of a transistor in common base configuration.

## 28.11) TRANSISTOR APPLICATIONS

The working principle of transistor is described in detail in the last lesson. Normally, the collector is reverse biased and no current flows in collector-emitter circuit. If we pass a very small current in the base circuit, a very large current starts flowing in the collector circuit. This property has made a transistor indispensable for vast electronic applications. But here we have discussed its applications as an amplifier, as a switch, and as an oscillator (frequency generator).

### 28.11.1 Transistor as an Amplifier

An electrical signal is voltage or current, which is coded with some useful information. For example, when we speak in front of a microphone, its diaphragm vibrates and induces a very small voltage in its coil, depending on the intensity of sound. This induced voltage appears as a weak signal and cannot operate a loudspeaker to reproduce sound. To make it intelligible, it is fed into a device called amplifier. The amplifier increases the level of input signal and gives out magnified output. If  $V_i$  is the input signal voltage fed to the amplifier and  $V_o$  denotes the amplified output, their ratio is called *voltage gain*.

$$A_v = \frac{V_o}{V_i} \quad (28.15)$$

Similarly, we can define the current gain and power gain as

$$A_i = \frac{i_o}{i_i} \quad (28.16)$$

$$A_p = \frac{P_o}{P_i} \quad (28.17)$$

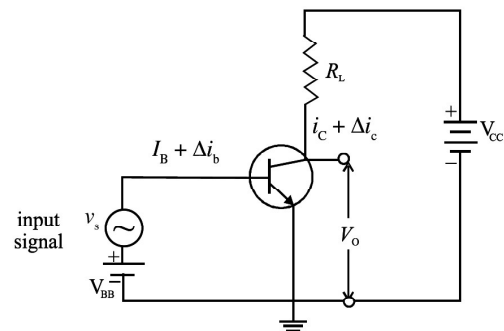


Fig. 28.36 : Basic amplifier circuit using n-p-n transistor in CE mode

The circuit for transistor as an amplifier is shown in Fig. 28.36.

Here an n-p-n transistor is used in CE mode. Its collector is reverse biased through the load resistance  $R_L$  by the battery  $V_{CC}$ . When a base current  $I_B$  flows, some collector current  $I_C$  will start flowing. On decreasing  $I_B$ , a stage will be reached when  $I_C$  becomes almost zero. This is the lower limit of variation of  $I_B$ . Similarly, on increasing  $I_B$  again, a stage of saturation is reached and  $I_C$  stops increasing. This corresponds to the upper limit of variation of  $I_B$ . For faithful amplification of input signal, a base current equal to the mean of these two limiting values of  $I_B$  is passed through the base by forward biasing it with battery  $V_{BB}$ . We can choose the operating point in the centre of linear operating range of the transistor. This is called biasing of the base. A signal source providing an input signal  $v_s$  is connected in series with  $V_{BB}$ .

Due to addition of oscillating signal voltage  $v_s$  to  $V_{BB}$ , the base current changes by an amount  $\Delta i_b$  around the dc biasing current  $I_B$ . The signal voltage is kept low so that the signal current  $\Delta i_b$  if added and subtracted from  $I_B$  does not cross the upper and lower limits of the base current variation. Otherwise, the transistor will go into cut off or saturation region and the amplified output will be highly distorted and noisy. Note that signal current

$$\Delta i_b = v_s / r_i \quad (28.18)$$

Where  $r_i$  is the input impedance. This change in base current  $\Delta i_b$  results in a large change in collector current, say  $\Delta i_c$  given by

$$\Delta i_c = \beta \Delta i_b = \beta v_s / r_i \quad (28.19)$$

Where  $\beta$  is the ac current amplification factor, equal to  $\Delta i_c / \Delta i_b$

From (Eqn. 28.19) we get

$$v_s = \Delta i_c \times r_i / \beta \quad (28.20)$$

By applying Kirchoff's law to the output circuit in Fig 28.36,

$$V_{CC} = V_{CE} + I_C R_L \quad (28.21)$$

This is the equation of load-line.

Differentiating Eqn. 28.21,

$$dV_{CC} = dV_{CE} + dI_C \times R_L \quad (28.22)$$

Since  $V_{CC}$  is constant,  $dV_{CC} = 0$ , then  $dV_{CE} = -dI_C \times R_L$

But  $dV_{CE}$  is the change in output  $\Delta v_o$  and  $dI_C$  in  $\Delta i_c$ . Therefore,

$$\Delta v_o = -\Delta i_c \times R_L$$

The voltage gain  $A_v$  of the amplifier is given by

$$\begin{aligned} A_v = v_o / v_s &= -(\Delta i_c \times R_L) / (\Delta i_c \times r_i / \beta) \\ &= -\beta \times R_L / r_i \end{aligned} \quad (28.23)$$

The ratio  $\beta/r_i$  is called **transconductance** of transistor and is denoted by  $g_m$ . Hence Eqn. (28.23) can be written as

$$A_v = v_o / v_s = -g_m \times R_L \quad (28.24)$$

The negative sign indicates that input and output are in opposite phase, i.e. they differ in phase by  $180^\circ$ . The power gain is given by

$$A_p = A_i \times A_v = \beta \times A_v \quad (28.25)$$

### John Bardeen (1908 – 1991)

John Bardeen is the only researcher in history of science who received two Nobel Prizes in Physics. He was born in Madison, Wisconsin USA, in a highly educated family. He was so bright a kid that his parents moved him from third grade to Junior high school. He did his graduation in Electrical Engineering. But, he also had to struggle for his career. After spending three years as geophysicist with Gulf Oil Company, he went to Princeton for his Ph.D. in Mathematical Physics. After a brief stint at Harvard and Minnesota and in Naval Ordnance Labs, he joined William Shockley's research group at Bell Laboratories. With Walter Brattain, he developed the first transistor for which Bardeen, Brattain and Shockley were conferred the 1956 Nobel Prize in Physics.



Bardeen shared his second Nobel in 1972 with Leon C Cooper and R Schieffer for their theoretical work on superconductivity.

### 28.11.2 Transistor as a switch

In day-to-day life, we use electrical switches to put the gadgets like lamps, fans, machines on or off manually. Note that the switch has two distinct states (on and off). In electronics, we come across situations where we need to apply an input to some device in the form of two distinct voltage levels. This is as if we were operating a switch. When switch is on, one voltage level is applied but when switch is off, the other one is applied. Typically, such voltage levels are used in computers, where digital signals are employed. This is done by using a transistor in the non-linear region of its operation.

In the transistor characteristics shown in Fig. 28.37 (a), we see two extreme regions: *cut-off region and saturation region*. The (jagged) region below the zero base ( $I_B = 0$ ) signifies the cut off region. The transistor does not conduct and entire supply voltage  $V_{CC}$  appears across the transistor between the collector and the emitter ( $V_{CE}$ ). That is, the output voltage at the collector is  $V_{CC}$ . When the base current  $I_B$  is greater than its saturation value, the transistor conducts fully and collector-emitter voltage  $V_{CE}$  is almost zero. In such a case, the output voltage obtained between collector and ground is zero and entire voltage drop appears across  $R_L$ . That is collector current is:

$$I_C = V_{CC} / R_L$$

Here, collector current  $= I_C = V_o/R$ . The control signal for switching the transistor ON or OFF states can be explained by  $V_{BB}$ . Fig. 28.37 (b) shows typical circuit diagram of transistor as a switch.

Input loop is:

$$I_B R_B + V_{BE} - V_{BB} = 0 \quad (28.26)$$

$$\text{If } V_{BB} = 0, \text{ then } I_B = -V_{BE} / R_B \quad (28.27)$$

For  $I_B$  less than zero, the transistor is cutoff and  $V_{CC}$  becomes  $V_o$ . By applying proper signals a transistor can be switched between saturation (on state) and cut off (off state). Hence, transistor acts as a switch.

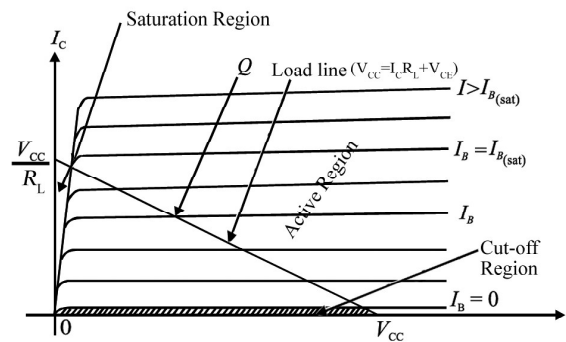


Fig. 28.37 (a) : Transistor output characteristics

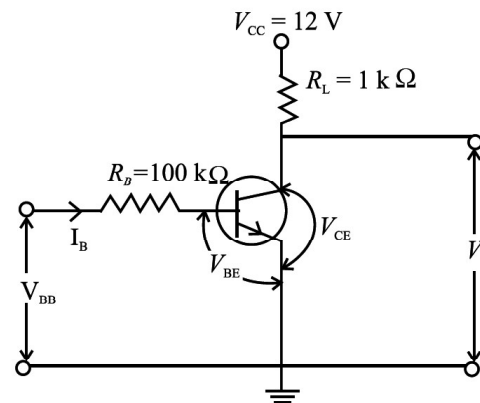


Fig. 28.37 (b) : Transistor as a switch

### 28.11.3 Transistor as an Oscillator

An electronic oscillator is a device which generates continuous electrical oscillations. In a simple oscillator circuit, a parallel LC-circuit is used as resonant circuit and an amplifier is used to feed energy to the resonant circuit. It can generate frequencies from audio to radio range depending on the choice of L and C.

We know that when a charged capacitor is connected across an inductor, the charge oscillates. But due to loss of energy by radiation and heating of wires, the energy is lost

and the amplitude of oscillations decays with time. To build a sinusoidal oscillator, where the oscillations are sustained (i.e. they do not decay), we need an amplifier with positive feedback. The basic idea is to feed a part of output signal in input signal. By adjusting the gain of the circuit and the phase of the feedback signal, energy dissipated in each cycle is replenished to get sustained oscillations of desired frequency. Schematically we can depict an oscillator to be made up of two main blocks: an amplifier with gain  $A$ , and a feedback circuit with feedback factor  $\beta$ , as shown in Fig. 28.38.

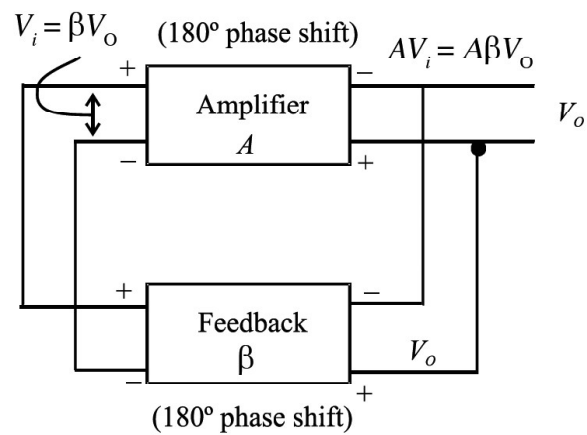


Fig. 28.38 : Schematic diagram of oscillator

In case,  $A\beta < 1$ ,  $V_o$  decreases continuously. On the other hand, if  $A\beta > 1$ ,  $V_o$  increases gradually. But if  $A\beta = 1$ , we get constant value of  $V_o$  leading to sustained oscillations.

Now, we consider a CE amplifier, discussed in the earlier section. It has  $180^\circ$  phase difference between the input and output, i.e. it has negative gain ( $-A$ ). To keep the total feedback gain  $A\beta = 1$ , we require that  $\beta$  is also negative; equal to  $-A^{-1}$ . That is, it is necessary to introduce a phase shift of  $180^\circ$  in the feedback circuit as well.

In Fig. 29.39, we have shown a circuit diagram of an oscillator using LC tank circuit and a transistor amplifier in CE mode. This is called Colpitts's Oscillator. In this circuit  $C_1$ ,  $C_2$  and  $L$  form the tank circuit. The oscillating current is generated in this circuit, which is at its resonant frequency. The output is obtained across  $C_1$ , the feedback is provided across  $C_2$  connected to the base of the transistor amplifier in CE mode. In this case  $180^\circ$  is introduced by the amplifier and another  $180^\circ$  phase shift is provided by the capacitor  $C_2$  which is connected between ground and other end of the inductor coil. Hence, the total loop gain is positive. When the gain of transistor amplifier is sufficiently large at the resonant frequency, we obtain sustained oscillations at the output.

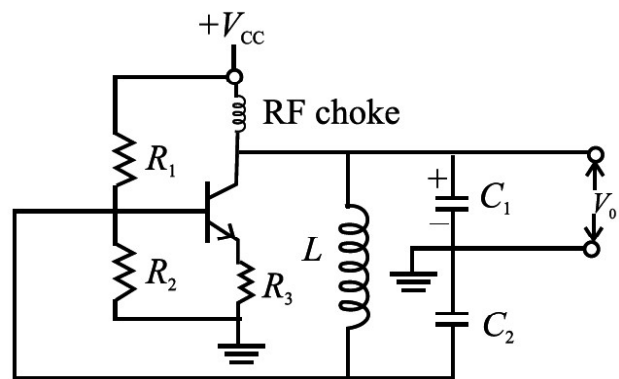


Fig. 29.39 : Colpitts's Oscillator

### Intext Questions 28.11

1. For a CE mode amplifier,  $V_i$  is 20 mV and  $V_o$  is one volt. Calculate voltage gain.
2. The  $P_o$  of an amplifier is 200 times that  $P_i$ . Calculate the power gain.
3. For CE amplifier,  $R_L = 2000$  ohms, and  $R_i = 500$  ohms and  $\beta = 50$ . Calculate voltage gain and power gain.



## 28.12) LOGIC GATES

In electronics, we come across mainly two types of waveforms. The information carried by these waveforms is called signal. When the signal takes any value within a range of amplitude at any instant of time, it is called a continuous signal. When the signal takes the value only at certain times, it is called a discrete signal. When the signal takes only particular finite number of amplitude values, it is called a **digital signal** (Fig. 28.40).

The digital signal varies in steps and typically has only two widely separated values '0' and '1'. These are called bits. Normally 0V corresponds to bit '0' and 5 V corresponds to bit '1'. Since the levels are so widely separated, any noise riding on the signal within the range of almost 2 V, [(0 V + 2 V) for level '0' and (5 V - 2 ;V) for level '1', does not affect the signal value, Hence these signals are immune to noise. The signals used in a computer are digital. The information is coded in the form of digital signals by a series of bits arranged in different order. Each bit is a pulse of fixed time duration.

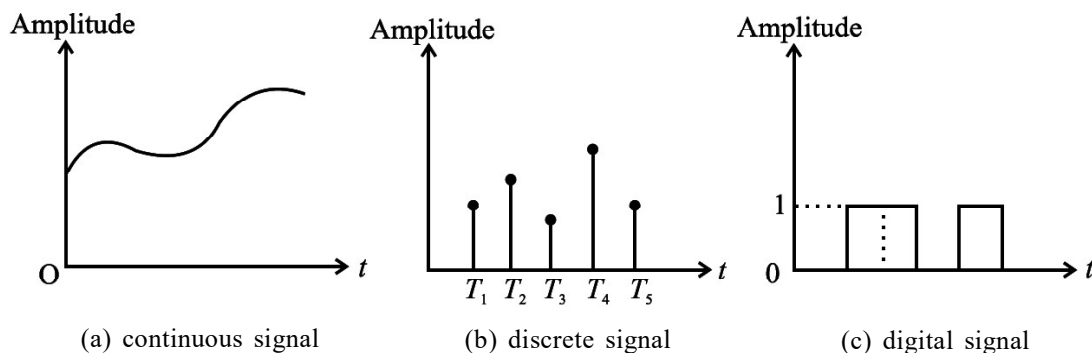


Fig. 28.40

Different mathematical operations can be performed on the digital signal. The mathematics governing these operations is called Boolean algebra. In Boolean algebra, the basic operations are addition and multiplication. If it is a digital data that takes value 0 or 1, the following identities hold:

$$A \times 0 = 0 \quad (28.28)$$

$$A + 1 = 1 \quad (28.29)$$

The circuits which perform these operations are called *logic gates*. Let us now learn about basic logic gates.

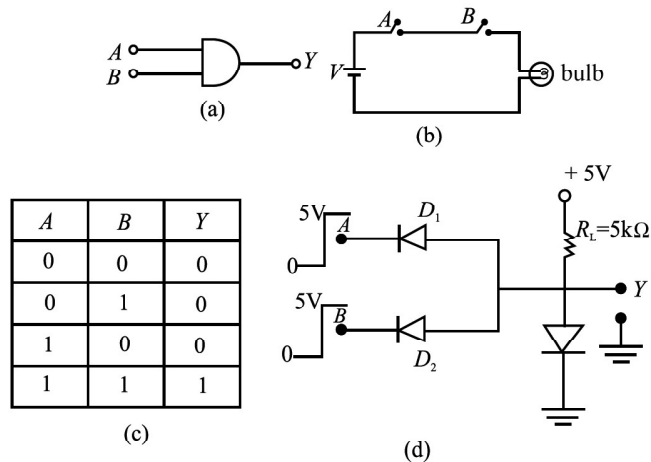
### 28.12.1 Basic Logic Gates

Logic gates are the devices which have one or more inputs and one output. They give different output when the input bits differ in their arrangement. The output produced by these gates follows the laws of Boolean logic. There are three basic types of logic gates:

1. AND Gate
2. OR Gate
3. NOT Gate

These gates perform multiplication, addition and inversion (negation) operations, respectively. Let us now learn the working of these logic gates.

1. **AND Gate :** An AND gate can have two or more inputs but only one output. The logic symbol of a two input AND gate is given Fig. 28.41 (a). We can understand the behavior of an AND gate by considering a number of electrical switches connected in series. For examples, switches A and B are two inputs of the gate and the bulb gives the output Y. The ON switch stands for logic input '1' and OFF switch stands for logic input '0'. In this case, the bulb will glow only if it is connected to the supply voltage. This will happen only if both A- and B- switches are simulta-neously ON (or '1'). The behavior of output Y at various values of A and B is shown in Table in Fig. 28.41 (c). This table is called *Truth Table*.



**Fig. 28.41 :** (a) symbol of AND gate (b) switch implementation of AND gate (c) truth table of AND gate (d) diode implementation of AND gate

The Boolean expression for the AND operation is represented as:

$$Y = A.B = AB = A \times B \text{ and read as A AND B.}$$

**Realization of AND Gate :** The logic gate realized by using diodes is called a DDL Gate (Diode–Diode Logic Gate). The diode implementation of a two-input AND gate is shown in Fig. 28.41 (d). The anodes of two diodes  $D_1$  and  $D_2$  connected in parallel are forward biased by a 5 V battery through a 5 k $\Omega$  resistance. The output is taken from the anode. Cathode wires A and B serve as input terminals. When either A or B or both the terminals are grounded, the respective diode will conduct and a potential drop will develop across the resistance and output will be 0.7 V, i.e. logic '0'. When both the terminals are connected to 5 V (i.e. for input 1, 1), neither of the diodes will conduct and output will be 5 V, i.e. logic '1'.

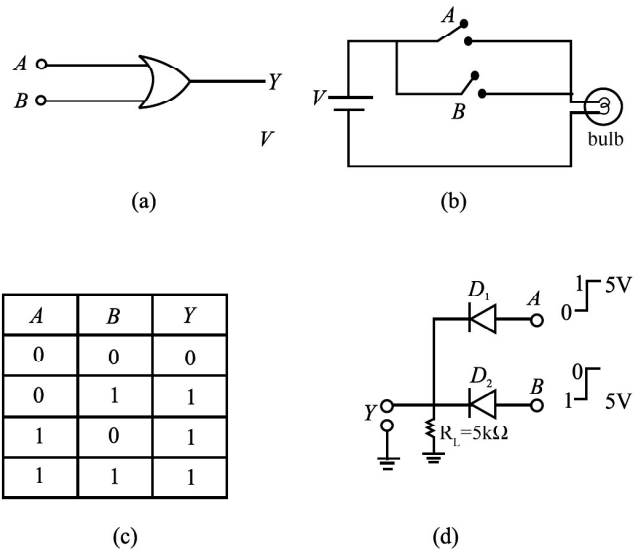
2. **OR Gate :** The OR gate can have two or more inputs and only one output. The logic symbol of a two input OR gate is given in Fig. 28.42 (a). We can explain the behavior of an OR gate with the help of a number of electrical switches connected in parallel. For a two input OR gate, two switches are connected, as shown in Fig. 28.42 (b). The switch A and B are the two inputs of the gate and the bulb gives output Y. The ON switch stands for logic input '1' and OFF switch stands for logic input '0'. The glowing bulb stands for logic output '1' and the non-glowing bulb for logic output '0'. In this case, when either A OR B or both the switches are ON, the supply voltage reaches the output and the bulb glows. The input-output correlation for an OR gate is shown in the Truth Table given in Fig. 29.42 (c).

The Boolean expression for an OR operation is represented as:

$$Y = A + B \text{ and read as A or B.}$$

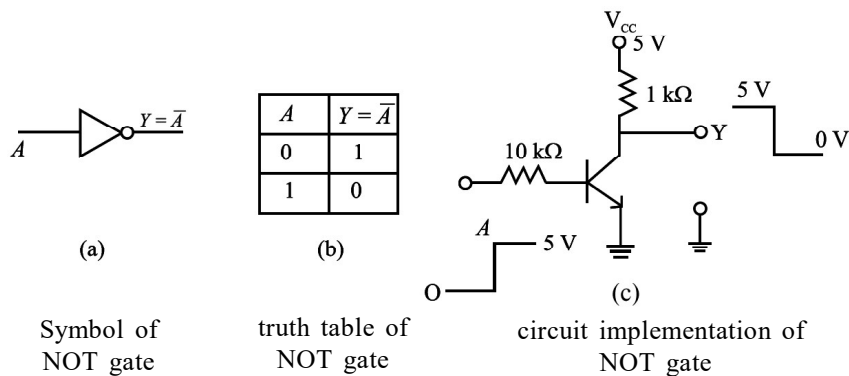
**Realization of OR Gate :**

The diode implementation of a two-input OR gate is shown in Fig. 28.42 (d). The cathodes of diodes  $D_1$  and  $D_2$  connected in parallel are grounded through a  $5\text{ k}\Omega$  resistance. The output is taken from the cathode and the two anode wires A and B serve as input terminals. When either A or B or both the terminals are connected to the positive terminal of the 5 V battery, the respective diode/diodes will conduct and potential at the output will be about 5 V i.e. logic '1'. When both the switches are open, output will be 0 V i.e. logic '0'.



**Fig. 28.42 :** (a) Symbol of OR gate (b) switch implementation of OR gate (c) truth table of OR gate (d) diode implementation of OR gate

3. **NOT Gate :** Another important gate used in digital signal handling is the **NOT** gate, which inverts the signal, i.e., if input is '1' then output of NOT gate is '0' and for '0' input, the output is '1'. The symbol for NOT gate is shown in Fig. 28.43 (a). The Truth Table of NOT gate is shown in Fig. 28.43 (b). The circuit to implement a NOT gate is identical to that used for a transistor as a switch. This is shown in Fig. 28.43 (c). When input A is at '0' level, transistor is off and the entire  $V_{CC}$  voltage (5V) appears at the output Y. When input A is '1' (5V), the transistor conducts and output voltage Y is '0'. The inversion operation is indicated by a bar on the top of the symbol of the input e.g. in the Truth Table we can write,  $Y = \text{NOT}(A) = \bar{A}$ .



**Fig. 28.43 :** NOT gate

So far we have discussed basic logic gates. You may now ask: Can we combine these to develop other logic gates? You will discover answer to this question in the following section.

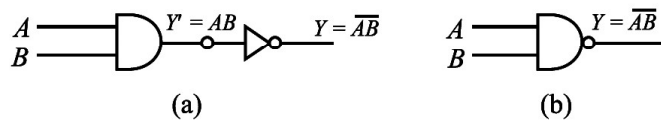
### 28.12.2 Combination Logic Gates

Two most important gates formed by combination of logic gates are :

- (1) NAND [NOT+AND]
- (2) NOR [NOT+OR] gates.

In digital electronics, a NAND gate or a NOR gate serve as building blocks because, use of multiple number of either of these gates allows us to obtain OR, AND and NOT gates. For this reason, these are called Universal gates. Let us now learn about combination logic gates.

1. **NAND Gate :** The NAND Gate is obtained by combining AND gate and NOT gate, as shown in Fig. 28.44 (a). Here the output  $Y'$  of AND gate is inverted by the NOT gate to get the final output  $Y$ . The logic symbol of a NAND gate is shown in Fig. 28.44 (b).



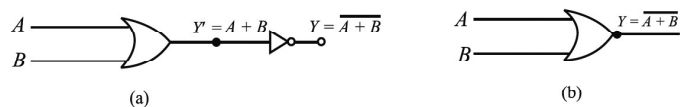
The Truth Table of a NAND gate is given in Fig. 28.44 (c). It can be obtained by inverting the output of an AND gate. The truth table of a NAND gate shows that it gives output '1' when at least one of the inputs is '0'. The Boolean expression of a NAND operation is represented as:

A	B	$Y' = AB$	$Y = \overline{AB}$
0	0	0	1
0	1	0	1
1	0	0	1
1	1	1	0

(c)  
**Fig. 28.44 :** (a) NAND as combination logic gate (b) symbol of NAND gate (c) truth table of a NAND gate

$$Y = \overline{A \cdot B}$$

2. **NOR Gate :** The NOR gate, obtained by combining an OR gate and NOT gate, is shown in Fig. 28.45 (a) Here the output of OR gate,  $Y'$ , is inverted by the NOT gate to get the final output  $Y$ . The logic symbol of a NOR gate is given in Fig. 28.45 (b).



A	B	$Y' = A+B$	$Y = \overline{A+B}$
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0

(c)  
**Fig. 28.45 :** (a) NOR as a combinational logic gate (b) symbol of NOR gate (c) truth table of NOR gate

The Truth Table of a NOR gate given in Fig. 28.45 (c), can be arrived at by inverting the output of an OR gate. The Truth Table of a NOR gate shows that it gives output '1' only when both the inputs are '0'. The Boolean expression for a NOR operation is represented as:

$$Y = \overline{A+B}$$

As mentioned earlier, the NAND and NOR gates are basic building blocks of all the logic gates. Let us now see, how we can obtain the three basic gates AND, OR and NOT by using NAND gates.

### 28.12.3 Realization of Basic Gates from NAND Gate

The NAND gate is considered to be the universal gate because all other gates can be realized by using this gate.

- (a) **Realization of a NOT gate :** If two input leads of a NAND gate are shorted together, as shown in Fig. 28.46, the resulting gate is a NOT gate. You can convince yourself about this by writing its truth table.

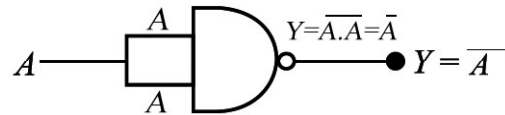
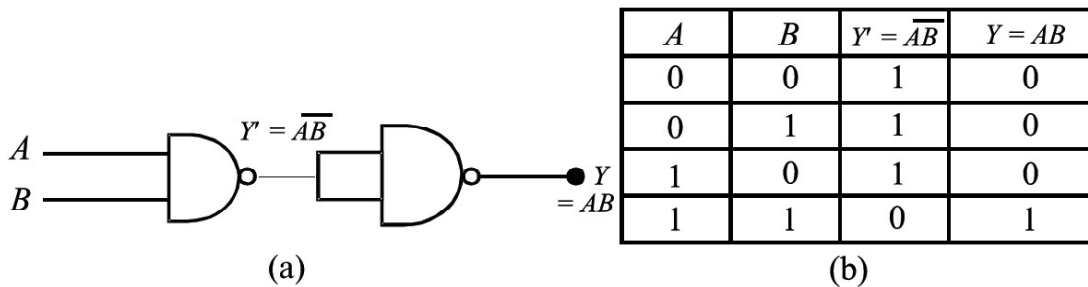


Fig. 28.46 : NAND gate as NOT gate

Here we have  $A = B$

- (b) **Realization of an AND gate :** The AND gate can be realized by using two NAND gates. The output of one NAND gate is inverted by the second NAND gate used as NOT gate as shown in Fig. 28.47 (a). The combination acts as an AND gate, as is clear from the Truth Table given in Fig. 28.47 (b).



NAND gates connected to implement AND gate      Truth table of AND gate using NAND gate

Fig. 28.47

- (c) **Realization of an OR gate :** The OR gate can be realized by using three NAND gates. Two NAND gates are connected as inverters and their outputs are fed to the two inputs of a NAND gate, as shown in Fig. 28.48. The combination acts as an OR gate.

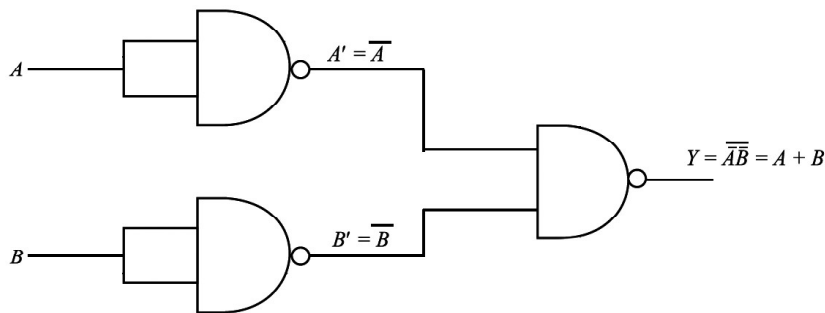


Fig. 28.48 : Three NAND gates connected as OR gate

**Intext Questions 28.12**

1. Complete the following table from Fig. 28.48 to prove that it is an OR gate.

A	B	A'	B'	Y
0	0	–	–	–
0	1	–	–	–
1	0	–	–	–
1	1	–	–	–

**28.13 INTEGRATED CIRCUITS**

The concept of fabricating an entire circuit on a small block or chip of a semiconducting technology has revolutionized the entire electronic technology. This small chip or circuit is called an integrated circuit (IC). In the conventional method, designing a circuit means by choosing the desired components like a diode, transistor, resistances (R), inductances (L), capacitors (C) etc., and connecting them by soldering wires. These circuits were found to be less reliable and less shockproof. Keeping this in view an entire circuit is fabricated on a small chip (IC). The most widely used chips are monolithic integrated circuits. The word monolithic is derived from the Greek word. Monos, means single and lithos mean stone. This means an entire circuit is formed on a single stone (or chip). The dimensions of the chip are usually  $1\text{ mm} \times 1\text{ mm}$  or even smaller than this. Figure 28.49 shows IC chip covered with protective plastic case. The connections can make by connecting the pins so that one can enable to use them for desired applications.

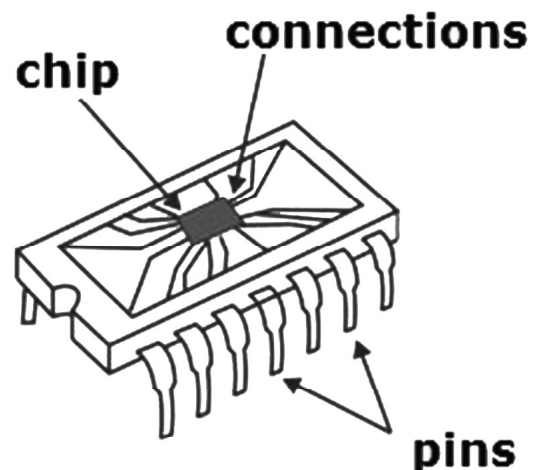


Fig. 28.49 : Monolithic IC in plastic case

Based on the input signals, IC are classified into two categories, namely Linear IC and Digital IC. Linear IC processes smoothly and continuously over a range of maximum and minimum values. The output is more or less directly proportional to the input value. In other words, the output signal varies linearly with the input. A good example of linear IC is the operational amplifier.

In the case of digital IC, the circuit contains logic gates. Depending upon the level of a number of circuitry operations or logic gates, these IC's are termed as: (i) SSI : small scale integration (logic gates  $< 10$ ) (ii) LSI : large scale integration (logic gates  $< 1000$ ) (iii) VLSI (very large scale integration, where logic gates  $> 1000$ ). The technology of fabrication is challenging large-scale production and has made it most inexpensive. Therefore they are future, faster and smaller devices.

**WHAT YOU HAVE LEARNT**

- Semiconductors are materials like silicon (Si) and germanium (Ge), which have conductivities midway between insulators and conductors.
- Semiconductors are of two types: Intrinsic (pure) and extrinsic (doped).
- Extrinsic semiconductors can be p-type (doped with 3rd group impurities) or n-type (doped with V<sup>th</sup> group impurities).
- A p-n junction diode consists of a n-type region and a p-type region, with terminals on each end.
- When a p-n junction is formed, diffusion of holes and electrons across the junction results in a depletion region which has no mobile charges.
- The ions in the region adjacent to the depletion region generate a potential difference across the junction.
- A forward biased p-n junction offers low resistance to flow of electrons.
- A reverse biased p-n junction diode offers high resistance to flow of current.
- A p-n junction allows current to flow in only one direction.
- A p-n junction diode can be used as a rectifier to convert ac into dc.
- A half-wave rectified dc contains more ac component than the full-wave rectified dc.
- A Zener diode stabilizes the output of a power supply.
- In a stabilizer, the Zener diode dissipates more power when the current taken by the load is less.
- A transistor consists of three separate regions (emitter, base and collector) and two junctions. Emitter is most heavily doped and base is the least doped. While collector has the largest size, base is the thinnest.
- Transistor can either be n-p-n type or p-n-p type.
- A transistor can be connected in any of the three configurations: common collector (CE), common base (CB) or common emitter (CE).
- The characteristics of a transistor vary according to the configuration of the transistor.
- CE configuration is preferred over other configurations as it provides high current gain and voltage gain.
- For amplification, a transistor needs input current.
- Transistor can be used as a switch by biasing it into saturation and cut-off regions.
- There are three basic logic gates: AND, OR and NOT.
- NAND gate is a universal gate because it can be used to implement other gates easily.
- Importance of integrated circuits.

### TERMINAL EXERCISE

1. Describe the most important characteristic of p-n junction diodes.
2. Explain the formation of depletion region in a p-n junction diode.
3. Which charge carriers conduct forward current in a p-n junction diode?
4. Differentiate between (i) Forward bias and reverse bias (ii) Avalanche and Zener breakdown
5. Explain the working of p-n-p and n-p-n transistors.
6. Define current gains  $\alpha$  and  $\beta$  of a transistor.
7. For  $\alpha = 0.998$ , calculate change in  $I_C$  if change in  $I_E$  is 4 mA.
8. Why the Peak Inverse Voltage (PIV) of a p-n junction diode in half-wave rectifier with filter capacitor is double of that without the capacitor?
9. Explain how a Zener diode helps to stabilize dc against load variation.
10. What should be the range of variation of amplitude of input signal for proper working of an amplifier?
11. Draw a circuit using diodes and transistors to implement a NOR gate.

### ANSWERS TO INTEXT QUESTIONS

#### 28.1

1. resistivity values
2. resistivity
3. more
4. eV
5. more

#### 28.2

1. zero
2. (ii)
3. impurity, doping
4. majority
5. lower



**28.3**

- (a) majority carriers (b) depletion region  
(c) 0.7, 0.3 (d) higher, electrons
- (a) (iii) (b) (iii) (c) (i)
- It is depleted of charge carriers and consists of ions only.

**28.4**

- Forward bias decreases the potential barrier height and width of the depletion layer.
- Reverse bias increases the potential barrier height and width of the depletion layer.
- (a) decreases (b) increases (c) increases, breakdown voltage
- (a) (iv) (b) (iii)
- Avalanche and Zener breakdown.

**28.5**

- (a) 0.7 V, 0.3 V (b) one (c) micro ampere
- (a) (iii) (b) (ii)

**28.6**

- (a) (ii) (b) (i) (c) (iv)

**28.7**

- See Fig. 28.16

In case of full wave rectifier, both diodes  $D_1$  and  $D_2$  charge C to maximum voltage of  $V_{\max}$  in alternate half cycles. Hence, the PIV of the diodes should be  $2 \times V_{\max}$ .

**28.8**

- $R_z = 100\Omega$ ,  $R_s = 100\Omega$  and  $R = R_z + R_s = 200\Omega$

Hence,

$$I = \frac{21}{200} = 0.105 \text{ A}$$

$$\text{and } V = IR = 0.105 \times 100 = 10.5 \text{ V}$$

2. (a) Zener (b) reverse (c) light sensitive  
 (d) group III-V (e) forward (f) emission  
 (g) recombination (h) electroluminescence (i) photovoltaic  
 (j) more, absorb

**28.9**

1. (a) (i) (b) (ii)  
 2. (a) Three, two (b) Base  
 (c) Most heavily, base (d) largest size, moderate  
 (e) Emitter-base, collector-base (f) npn, pnp

**28.10**

1. (a) input characteristic (b) output characteristics  
 (c) Collector, emitter (d) base and emitter, base and collector

**28.11**

1.  $A_v = V_o / V_i = 1 \text{ V} / 20 \text{ mV} = 50$   
 2.  $A_p = P_o / P_i = 200$   
 3.  $A_v = \frac{(\beta \times R_L)}{r_i} = \frac{50 \times 2000 \Omega}{500 \Omega} = 200$   
 $A_p = \beta A_v = 50 \times 200 = 10000$

**28.12**

A	B	A'	B'	Y
0	0	1	1	0
0	1	1	0	1
1	0	0	1	1
1	1	0	0	1

**ANSWERS TO TERMINAL EXERCISE**

7. 3.992 mA



# BASICS OF COMMUNICATION

## INTRODUCTION

In the world of animals, communication is made by mechanical, audio and chemical signals. You might have observed how sparrows begin to chirp loudly on seeing an intruder, who can put their life in danger! However, human beings are blessed with very strong means of communication. We can express what we see, think and feel whatever is happening around us. We use sound (an audible range, 20 Hz – 20 kHz) and light (in visible range, 4000 Å – 7000Å) for communication. Communication entails transmitting and receiving information from one individual / place to another. *In other words, the receiver and sender should understand a common language.*

Prior to the written alphabet, in the first era, the mode of communication was oral. The second era of communication began with the invention of printing press. Invention of the telegraph in the early nineteenth century marked as the beginning of the third stage. Revolutionary technological developments enabled us for efficient and faithful transfer of information. Using tools and techniques such as telegraph, fax, telephone, radio, mobiles, satellites, computers etc., it is possible to communicate over long distances. The oceans and mountain ranges have no longer posed any problem. The constraint of time and distance is also non-existent here. During recent pandemic or post-pandemic, on-line learning, (education), publishing (research), banking (business) which are now becoming routine activities. In fact, communication has opened a new horizon.

Modern communication system has classified into the following three categories.

- One to one (point-to-point) communication.
- One to many (broadcast) communication.
- Many to many (mobile conference call) communication.

In communication system, the information is in the form of electrical signal (voltage or current) spread over a range of frequency (called the signal bandwidth). Noise is also added to the signal and tries to obscure the desired information.

## OBJECTIVES

After studying this lesson, you should be able to

- *understand the list the components used for communication system;*
- *explain the terms like analog, digital, modulation and demodulation signals etc.;*
- *describe how electromagnetic waves act as carriers of information;*
- *describe the processes of modulation and demodulation.*

## 29.1 COMPONENTS OF A COMMUNICATION SYSTEM

Fig. 29.1 shows building blocks of a typical communication system. The essential elements of a communication system are:

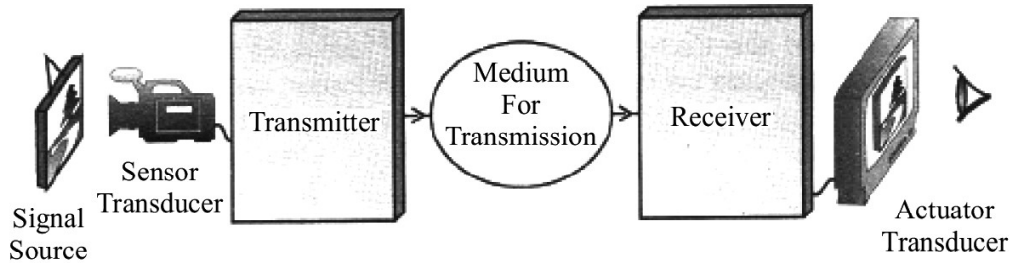
- a source of signal, a sensor transducer and a **transmitter**, which launches the signal carrying information,
- an intervening **medium/channel** to guide and carry the signal over long distances
- signal **receiver** (an actuator transducer) to intercept the signal and retrieve the information.

The most commonly used signals in communication are either audio or visual. These are characterized by amplitude, frequency, phase and polarization. For sound signals, the audible range is 20 Hz – 20 kHz. In case of normal telephone, the range is limited to 4 kHz only. Another noteworthy aspect is that the policemen (or security personal) carrying Walkie-Talkie sets to monitor the movement of dignitaries or public rallies. The range of such sets is limited to 1 kHz.

Some of the commonly used terms for communication are:

- (i) **Transmitter and receiver** : An input signal (bearing information) is transmitted to a distant point by a transmitter. A **receiver** intercepts such signals and transforms them in such a way that the information hidden therein can be converted into usable form.
- (ii) **Signal** : A clerical form of information which is suitable for transmission is called signal. A signal can be either analog or digital. Analog signals are continuous-variations of voltage or current. Analog signals are single-valued functions of time. Digital signals are those which can take only discrete stepwise values.
- (iii) **Noise** : The unwanted signals that tend to disturb the transmission signals in communication system is called noise.
- (iv) **Attenuation** : The loss of strength of a signal while propagating through a medium is called attenuation.
- (v) An **antenna** or **aerial** is essentially systems of conductors, which effectively radiates and absorb electromagnetic waves. The antenna can be in the form of a long, stiff wire (as AM/FM radio antennas on most cars) or a huge dish (for far away satellites).
- (vi) **Sensitivity** signifies the minimum input voltage required to produce a standard output signal voltage. The greater the amplification of the receiver, the greater is its sensitivity. Signal to noise ratio, abbreviated as the **S/N ratio**, plays an important role in determining the sensitivity of a receiver. If any external noise enters the systems, it must be filtered out using some signal processing technique.
- (vii) **Selectivity** is the capability of a receiver to differentiate between a desired signal of a particular frequency and all other unwanted signals of nearby frequencies. Selectivity depends on the sharpness of the resonance curves of the tuner circuits used in the receiver.

- (viii) **Fidelity** represents the variation of the output of a receiver with the modulation frequency and denotes the ability of the receiver to reproduce the waveform of the modulating signal.



**Fig. 29.1** : A schematic arrangement for the communication system.

For transmission the input signal (say radio signal) is usually in the form of voice or music and the transmitter transforms it (by a process called **modulation**, which you will learn in the end of this chapter) into electrical signal (by superposing over electromagnetic waves in the frequency range 30 kHz – 300 MHz). These radio signals are broadcast by means of aerials or antennas either in all directions or in some specified direction. In a radio transmitter the antenna launches the radio waves into space. In a receiver, the idea is to pick up maximum transmitted power and supply it to the tuner. The **optimum size** of a radio antenna is related to the frequency of signal. Antenna tries to transmit or receive. The size of these conductors has to be comparable to the wavelength  $\lambda$  of the signal (at least  $\lambda/4$  in dimension), so that they can detect the time-variation of the signal properly. If the signals picked up by the receiving antenna are extremely weak, then the signals are amplified before being analyzed. In brief, a signal is communicated from the transmitter to the receiver through a medium. The carrier is in the form of a wave and for sound as well as electromagnetic waves. So medium supports the principle of superposition. The aim of the communication system is faithful reproduction of transmitted signal. A circuit designer makes every effort to ensure best possible reproduction at the receiving end.

### Intext Questions 29.1

1. What is the frequency range of radio waves?
2. How do you determine the optimum size of a radio antenna?

## 29.2 TYPES OF SIGNALS

Signals are classified on the basis of their origin and nature, accordingly we have:

- continuous time (analog) and discrete time (digital) signals;
- coded and un-coded signals
- periodic and aperiodic signals
- energy and power signals
- deterministic and random signals.

We will consider here only analog and digital systems. In modern electronic communication systems, the signals are converted into discrete form. They have finite values of amplitude at different instances of-time and zero [see Fig. 29.2 (b) and (c)] form. From the Fig. 29.2, you will note that the waveforms used to represent correspond to a particular frequency and are periodic; while one of these is sinusoidal, another is pulsed. In fact these may be viewed as a sub-class of sine and square waveforms. Information can be packaged in both analog (continuous) and digital (discrete) forms. Speech, for example, is an analog signal which varies continuously with time. In contrast, computer files consist of a symbolic “discrete-time” digital signal.

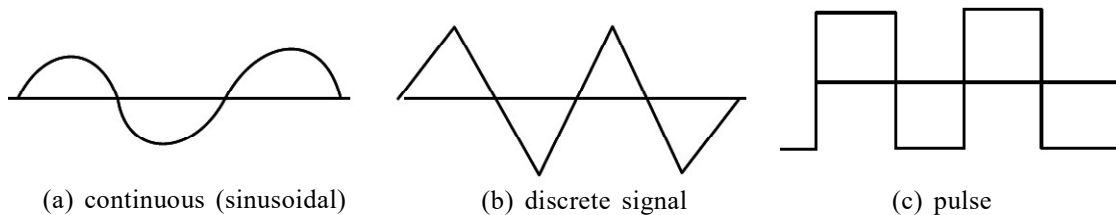


Fig. 29.2 : Types of signals

In the digital format, signals are in the form of a string of **bits** (abbreviated from binary digits), each bit being either ‘ON’ or ‘OFF’ (1 or 0). The **binary** system refers to a number system which uses only two digits, 1 and 0. We can convert all information-bearing signals into discrete-time and amplitude-quantized digital signals. In a compact disc (CD), the audio is stored in the form of digital signals, just as a digital video disc (DVD) stores the video digitally. Analog systems are in general, less expensive than digital systems. But digital systems are more efficient and give better performance (less error and noise), and greater flexibility. Interestingly, digital as well as analog transmission accomplished using analog signals, like voltages in Ethernet (an example of wire-line communication) and electromagnetic radiation in cellular phone (wireless communication). Transmitting the signal to large area and to lower transmitting power is the main key idea for the mobile communication. The hexagon concept designed by Subasinghe-Dias (1992) is used to serve small area, called cell, made a hall mark approach to the modern mobile communication system.

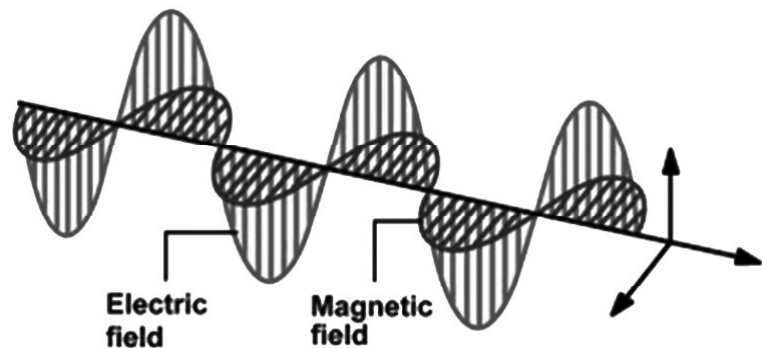
The most crucial parameter in communication systems is the signal bandwidth, which refers to the frequency range in which the signal varies. Analog bandwidth measures the range of spectrum and bandwidth gives the quantity of information contained in a digital signal. For this reason, analog bandwidth is expressed in terms of frequency (Hz) and the digital bandwidth is expressed in terms of bits per second (bps). The frequency range of some audio signals and their bandwidths are given in Table-29.1. Note that human speech has bandwidth of nearly four kilo hertz. The bandwidth is about 10 kHz in amplitude modulated (AM) radio transmission and 15 kHz in frequency modulated (FM) transmission. However, the quality of signal received from FM broadcast is significantly better than that from AM. The compact discs have bandwidth of 20 kHz. The bandwidth of a video signal is about 4.2 MHz and television broadcast channel has bandwidth of 6 MHz. The bandwidth of a typical modem (a device used for communication of digital signals over analog telephone lines) systems are 32 kbps, 64 kbps or 128 kbps respectively.

**Table - 29.1 : Frequency and bandwidths for known systems**

Source	Frequency range (Hz)	Bandwidth (kHz)
Guitar	82–880	... 0.8
Violin	196–2794	... 2.6
Vowels (a,e,i,o,u) consonants	250–5000	... 4
Telephone signal	200–3200	... 3

Wireless communication was originated by Dr. J.C. Bose and G. Marconi. Since then, technology has moved a long way. At earlier stage, radio communication was made via amplitude modulated (AM) transmission. Subsequently, it gave a way to frequency modulated (FM) transmission. Presently, radio signals are transmitting by means of satellite-communication. But for simplicity, we here confine ourselves to using analog and digital signals, AM and FM modulation and demodulation processes. Since the digital system of communication is more efficient and noise-free, we intend to convert an analog signal to its digital counterpart. This is done using sampling technique. A detailed aspect of this would be discussed in next section.

In communication, different ways are there to transport the electrical signal from the transmitter to the receiver. However, the current passes through a metal conductor in the form of current signal or voltage drop, through air in the form of electromagnetic radiation. The term electromagnetic suggests that the waves consist of electric and magnetic fields, which are inseparable. The e. m. wave propagation was given by Maxwell. According to him, an electric field varying in time produces a space-time varying magnetic field, which, in turn, produces electric field. The pictorial representation of a plane electromagnetic wave is shown in Fig. 29.3.

**Fig. 29.3 :** Propagation of e .m. waves

Mathematically, electric and magnetic field vectors can be represented as:  $E = E_0 \sin(kz - \omega t)$  and  $H = H_0 \sin(kz - \omega t)$ . The direct experimental evidence for the existence of e. m. waves came in 1888 through a series of brilliant experiments initiated by Hertz. In 1895 Indian physicist Jagadish Chandra Bose produced waves of wavelength in the range 25 mm to 5 m and demonstrated the possibility of radio transmission. A detailed e.m. spectrum and other wave spectrum range is given in the Fig. 29.4. This work was put to practical use by Guglielmo Marconi who, succeeded in transmitting the electromagnetic waves across the Atlantic Ocean. This marked the beginning of the new era of communication using e.m. waves. Marconi along with Carl Ferdinand Braun, received the 1909 Nobel Prize in physics for their work on wireless telegraphy.

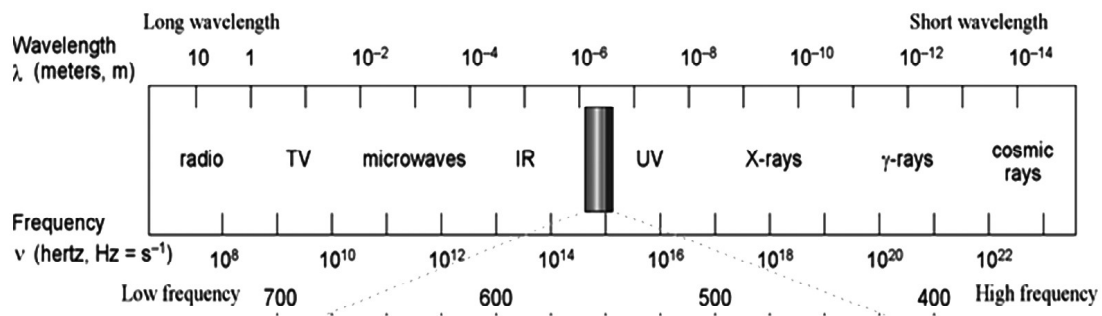


Fig. 29.4 : The electromagnetic and other wavelength spectrum

In a communication system, a transmitter radiates electromagnetic waves with the help of an antenna. These waves propagate in the space and captured by the receiver. At the receiver, another antenna extracts energy (in formation) from the electromagnetic waves. Now we use radio waves for different purposes television (TV) broadcasts, AM (amplitude modulated) and FM (frequency modulated) radio broadcasts, police and fire radios, satellite-TV transmissions, cell phone conversations, and so on. Each such signal uses a different frequency, and that is how they are all separated. In Table-29.2, we have listed internationally accepted electromagnetic spectrum relevant for radio and TV broadcast, popular brand names, and their application.

Table - 29.2 : Radio frequency bands

Band	Frequency Range	Wavelength Range	Application
Extremely Low Frequency (ELF)	< 30 Hz	> 10,000 km	under water communications
Very Low Frequency (VLF)	3 - 30 kHz	100 – 10 km	SONAR
Low Frequency (LF)	30 - 300 kHz	10 – 1 km	marine navigater
Medium Frequency (MF)	300 kHz - 3 MHz	1 km – 100 m	medium wave radio
High Frequency (HF)	3 - 30 MHz	100 – 10 m	short wave radio,
Very High Frequency (VHF)	30 – 300 MHz	10 – 1 m	FM radio, magnetic resonance imaging
Ultra High Frequency (UHF)	300 MHz – 3 GHz	1 m – 10 cm	commercial, TV, radio, radar wireless communication
Super High Frequency (SHF)	3 – 30 GHz	10 – 1 cm	modern communication satellite communication, cellular mobile, commercial TV

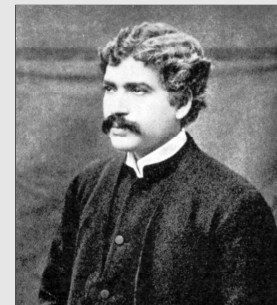


Based on the broadcast on bands, waves are popularly classified as:

- (i) **Medium wave (530-1700 kHz)** has been most commonly used for commercial AM radio broadcasting.
- (ii) **Long wave (144-351 kHz)** is used everywhere except in North and South Americas, where this band is reserved for aeronautical navigation. For long- and medium-wave bands, the wavelength is long enough that the wave diffracts around the curve of the earth by ground wave propagation, giving AM radio a long range, particularly at night.
- (iii) **Short wave (3-30 M Hz)** is used by radio services intended to be heard at great distances away from the transmitting station; the far range of short wave broadcasts comes at the expense of lower audio fidelity. The mode of propagation for short wave is ionospheric.

### Jagadish Chandra Bose (1858 – 1937)

Jagadish Chandra Bose, after completing his school education in India, went to England in 1880 to study medicine at the University of London. Within a year, he took up a scholarship in Cambridge to study Natural Science at Christ's College – one of his lecturers at Cambridge; Professor Rayleigh had a profound influence on him. In 1884 Bose was awarded B.A degree by Cambridge University and B.Sc degree by London University. Bose then returned to India and took teaching assignment as officiating professor of physics at the Presidency College in Calcutta (now Kolkata). Many of his students at the Presidency College were destined to become famous in their own right. Satyendra Nath Bose who became well known for his pioneering work on Bose-Einstein statistics and M.N. Saha who gave revolutionary theory of thermal ionization, which enabled physicists to classify the stars into a few groups.



In 1894, J.C. Bose converted a small enclosure adjoining a bathroom in the Presidency College into a laboratory. He carried out experiments involving refraction, diffraction and polarization. To receive the radiation, he used a variety of junctions connected to a highly sensitive galvanometer. He developed the use of *galena* crystals for making receivers, both for short wavelength radio waves and for white and ultraviolet light. In 1895, Bose gave his first public demonstration of radio transmission, using these electromagnetic waves to ring a bell remotely and to explode some gunpowder. He invited by Lord Rayleigh, to give a lecture in 1897. Bose reported on his microwave (2.5 cm to 5 mm) experiments to the Royal Institution and other societies in England. But Nobel Prize alluded him probably for want of vivid practical application of this work by him. By the end of the 19<sup>th</sup> century, the interests of Bose turned to response phenomena in plants. He retired from the Presidency College in 1915, and was appointed Professor Emeritus. Two years later the Bose Institute was founded in Kolkata. Bose was elected a Fellow of the Royal Society in 1920.

**Intext Questions 29.2**

1. What is an electromagnetic wave?
2. Calculate the wavelength of a radio wave of frequency of 30 MHz propagating in space.
3. What is the frequency range of visible light?

**29.3) SAMPLING THEOREM**

In the preceding section, you may recall that an analog signal has an infinite number of very precise values in a certain time interval. Since it is impossible to count and store its values at infinitesimally close instants of time with infinite precision, we devise a practical way of picking a good digital approximation. The first step in this process is **sampling**. To sample a signal, we note its values at regular intervals of time (the rate at which the samples are taken is called sampling rate).

**Sampling theorem** states that, an analog signal is completely described by its samples, taken at equal time intervals  $T_s$ , if and only if the sampling frequency  $f_s = 1 / T_s$  is greater than or equal to twice the maximum frequency component (i.e., the bandwidth) of the analog signal. The equality defines what is called the **Nyquist rate**. Thus by sampling an analog signal, the signal is converted (without any loss of information) into an amplitude continuous time-discrete signal, which in turn can be converted a quantized wave into a signal discrete in both amplitude and time. This means that the infinite precision values of amplitudes are converted to values which can be stored digitally.

As mentioned earlier, the three different types of analog modulation are **Amplitude Modulation (AM)**, **Frequency Modulation (FM)** and **Phase Modulation (PM)** respectively. For pulsed carrier waves, pulse code modulation (PCM) is preferred. A detailed aspect of this would be discussed in the following section.

**29.4) MODULATION ANALOG (AM and FM) AND DIGITAL (PCM)**

The processing of signal for transmission is called *modulation*. The information-bearing signals are audio signals of frequency, less than 20 kHz. For small distances, we can communicate by simple a direct link. But it is not practical to transmit such signals to long distances, because of the following two reasons:

- 1) Signal should have an antenna/aerial of size and comparable to the wavelength of the signal so that the time variation of the signal is properly sensed by the antenna. It should remember here that for low-frequency (or long-wavelength) signals, the antenna size should be large.
- 2) If the power carried by low frequency signals is small then waves cannot transmit for large distance. This is mainly attributed to the continuous decline (or automation) due to absorption/radiation loss. In other words, high frequency signal are used for long distance transmission. But these signals do not carry any useful information. For better understanding, a simple event is discussed in the following paragraph.

Let us say, our Indian army spots advancing enemy forces. In order to save the post, army needs reinforcement from the base camp. But by the time an army jawan goes, conveys the message and the reinforcement reaches, the post would have fallen. Therefore, he wants a carrier; say a horse, which can run fast. But the horse cannot deliver the message. Therefore, the way out is: put the jawan on the horseback; let the horse run and jawan convey the message. For signal transmission, audio signal acts as jawan and high (radio) frequency acts as the horse (carrier). In other words by super imposing a low frequency signal on a high frequency carrier wave, one can process a signal and make it suitable for transmission. We convert the original signal into an electrical signal, called the *base band signal* using a signal generator. The carrier wave can be continuous or pulsed. Since a sinusoidal wave, is characterized by amplitude, frequency and phase, it is possible to moderate either of these physical parameters. This is known as analog modulation and they are two types: **Amplitude Modulation (AM)** and **Frequency Modulation (FM)**.

**Amplitude modulation:** The amplitude of a high-frequency carrier wave [Fig. 29.5 (a)] is modified in accordance with the strength of a low-frequency audio or video modulating signal [Fig. 29.5 (b)]. If the amplitude of the modulating wave increases, then the amplitude of the modulated carrier also increases. The envelope of the modulated wave takes the form depending on the amplitude and frequency of modulating signal [Fig. 29.5 (c)].

Let the instantaneous amplitudes of audio signal and carrier wave are:

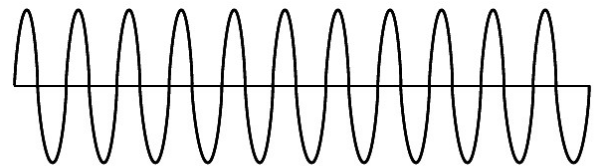
$$v_a(t) = v_{ao} \sin \omega_a t \quad (29.1 a)$$

and

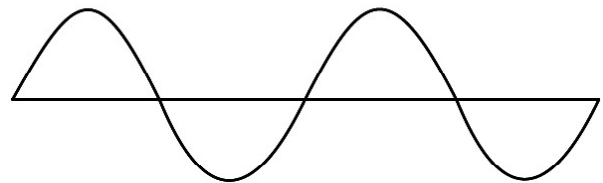
$$v_c(t) = v_{co} \sin \omega_c t \quad (29.1 b)$$

Here,  $\omega_a$  and  $\omega_c$  are the angular frequencies and  $v_{ao}$  and  $v_{co}$  denote velocities of audio and carrier waves, respectively. Amplitude modulation signal is superimposed on the carrier wave, so that the amplitude of the resultant modulated wave can be expressed as:

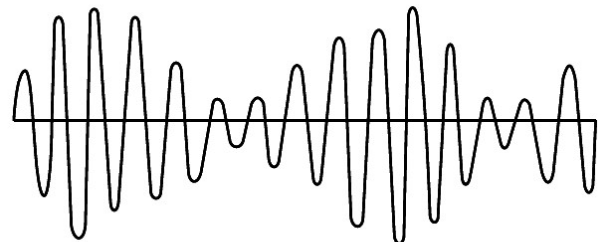
$$\begin{aligned} A(t) &= v_{co} + v_a(t) = v_{co} + v_{ao} \sin \omega_a t \\ &= v_{co} \left[ 1 + \frac{v_{ao}}{v_{co}} \sin \omega_a t \right] \end{aligned} \quad (29.2)$$



(a) a sinusoidal carrier wave of high frequency



(b) a modulating signal (message or information signal) of low frequency,



(c) amplitude modulated carrier wave

**Fig. 29.5 :** Modulation of a carrier wave by a modulating signal

Hence, the modulated wave can be expressed as:

$$v_c^{\text{mod}}(t) = A \sin \omega_c t = v_{co} \left[ 1 + \frac{v_{ao}}{v_{co}} \sin \omega_a t \right] \sin \omega_c t \quad (29.3)$$

From Eqn. (29.3) we note that the instantaneous amplitude of the modulated wave is determined by the amplitude and frequency of the analog audio signal. The ratio  $v_{ao} / v_{co}$  gives us a measure of the extent to which carrier amplitude is varied by the analog modulating signal and is known as amplitude modulation index. We will denote it by  $m_a$ . In terms of modulation index, we can rewrite Eqn. (29.3) as:

$$\begin{aligned} v_c^{\text{mod}} &= v_{co} (1 + m_a \sin \omega_a t) \sin \omega_c t \\ &= v_{co} \sin \omega_c t + v_{co} m_a \sin \omega_a t \sin \omega_c t \\ &= v_{co} \sin \omega_c t + \frac{v_{co} m_a}{2} \cos(\omega_c - \omega_a)t - \frac{v_{co} m_a}{2} \cos(\omega_c + \omega_a)t \end{aligned} \quad (29.4)$$

From the above equation it is evident that:

- (i) The modulated wave [shown in Fig. 29.5 (c)] has three components. The first term represents carrier wave. The second term frequency is lower than that of the carrier wave and constitutes at the lower side band. The third term with frequency higher than the carrier wave is the upper side band.
- (ii) The frequency of the modulating signal is not directly contained in the amplitude modulated wave.

It is important to note the most efficient information transfer takes place when maximum power transmitted by the communication system is contained in the side bands. The block diagram of a basic analog AM-transmitter is shown in Fig. 29.6 (a). The oscillator provides a fixed frequency and the power amplifier modulates the signal. In the AM broadcast transmitter, Hartly-Colpitts oscillators are used [Fig. 29.6 (b)] . Moreover a better amplifier is introduced between the master oscillators in the remaining circuit. To enhance the frequency and amplitude of the signal frequency multiplier and driver amplifiers are added before modulating the signal in a power amplifier.

For any broadcast, the maximum power of radio transmitter is in the range 500 W to 50 kW. Every broadcaster, wave should allocate a definite frequency and wave

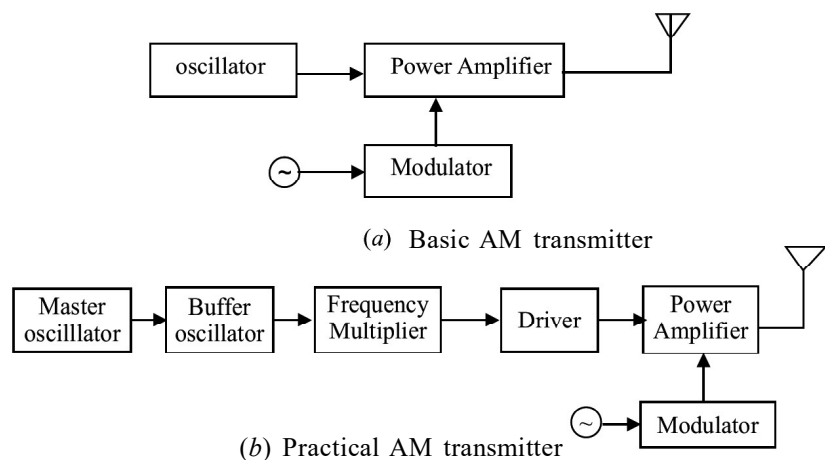


Fig. 29.6 : Block diagram of AM transmitter

should not interfere with other signals. To ensure this, undesirable frequencies are filtered out by using coupling circuits.

The radio communication of India, over the past 50 years, is being used medium wave (530 – 1700 kHz) and short wave (3 – 30 MHz) for analog AM. It continues to have the widest spread, through analog FM broadcast and now being preferred because of better quantity. Moreover, radio waves are now comparatively free and private broadcasters are also entering the field in a big way. FM radio stations are also being created by educational institutions for education as well as empowerment of rural youth and homemakers. In TV transmission, audio is frequency modulated, whereas the video (picture) the amplitude is modulated.

**Frequency modulation:** Here, the amplitude of the carrier wave remains constant and its frequency is continuously varied in accordance with the instantaneous amplitude of the audio/video signal. If the amplitude of the modulating signal voltage is large, then the carrier frequency goes up. On the other hand, if the amplitude of the modulating signal is low, then the carrier frequency goes down. The frequency of the FM wave will vary from a minimum to a maximum, corresponding to the minimum and maximum values of the amplitude of the modulating signal (see Fig. 29.7).

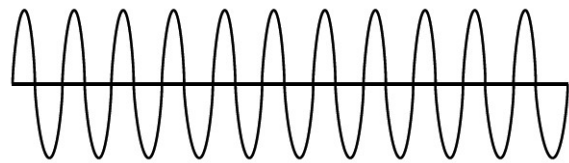


Fig. 29.7 : Frequency modulated carrier wave

An FM Transmitter essentially contains an oscillator, whose frequency of the carrier is varied depending on the input audio signal. It is usually accomplished by varying capacitance in an LC oscillator or by changing the charging current applied to a capacitor, (for example, by the use of a reverse biased diode, since the capacitance of such a diode varies with applied voltage). After enhancing the power of the modulated signal, it is fed to the transmission antenna. Low-frequency radio broadcast stations use this type of frequency modulation, since it is a simple, robust method.

**Phase modulation** involves changing the phase angle of the carrier signal in accordance with the modulating frequency. Analog pulse modulation is either amplitude modulated or time modulated. Similarly, digital pulse modulation is of two types: (i) pulse code modulation and (ii) pulse delta modulation.

Here, in the pulse code modulation (PCM), the modulating signal is first sampled, and the magnitude (with respect to a fixed reference) of amplitude of each sample is quantized. It is a digital representation of an analog signal where the magnitude of the signal is sampled regularly at uniform intervals of duration  $T_s$ . The binary code is transmitted usually by modulating an analog current in a transmission medium such as a landline. PCM is used to convert analog signals into digital signals. The pulse code modulation is used in digital telephone systems and digital audio recording on compact discs.

## 29.5) DEMODULATION

The modulated signal carrying the information, once radiated by the antenna, travels in space. Since there are so many transmitting stations and thousands of signals reach our

antenna. We have to choose the desired signal and decouple the carrier wave and the modulating signal. This process is known as *demodulation*. In other words, in demodulation process, we eliminate the radio frequency carrier and separate (filter) the modulation signal. For performing amplitude demodulation, the simple circuit shown in Fig. 29.8 (a) is used. The modulated wave received by the antenna is fed into a resonant circuit. The various RF voltages induced in the receiver aerial, only those having the same frequency as the resonant frequency of LC circuit are turned in due to electromagnetic induction between coils  $L_1$  and  $L$ . By varying  $C$ , the resonant frequency of the LC circuit can be varied and hence RF signal of any desired frequency can be turned in. This input signal is rectified by the diode and passed on to the low-pass filter  $RC_1$ .

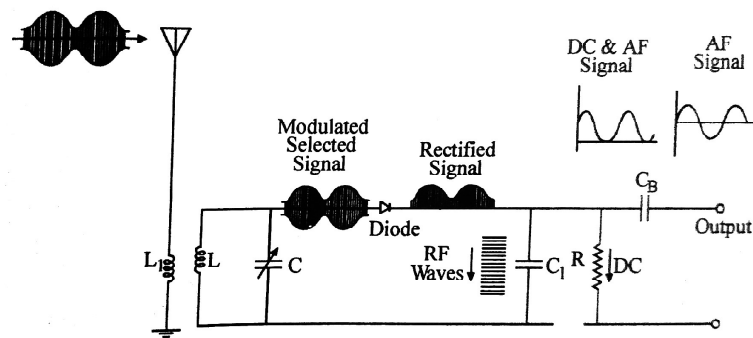


Fig. 29.8 (a) : Circuit diagram of demodulated wave

The modulated wave received by the antenna is fed into a resonant circuit. The various RF voltages induced in the receiver aerial, only those having the same frequency as the resonant frequency of LC circuit are turned in due to electromagnetic induction between coils  $L_1$  and  $L$ . By varying  $C$ , the resonant frequency of the LC circuit can be varied and hence RF signal of any desired frequency can be turned in. This input signal is rectified by the diode and passed on to the low-pass filter  $RC_1$ .

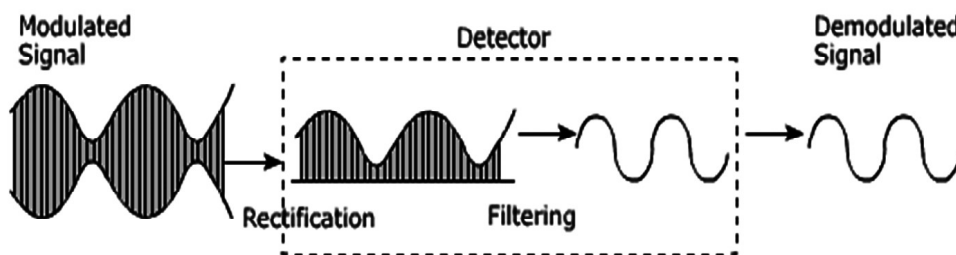


Fig. 29.8 : (b) input and output waves

The RF carrier wave is filtered out by low-reactance capacitor  $C_1$  which is called RF filter capacitor or RF by-pass capacitor. The dc component of the remaining signal is shunted out through  $R$  because it cannot pass through blocking capacitor  $C_B$ . But the low, frequency AF signal can easily get through  $C_B$  and becomes available across the output. When passed through a suitable device, say, a headphone, the original sound can be heard. Fig. 29.8 (b) shows modulated and demodulated waves.

## 29.6) SPACE COMMUNICATION

The wireless communication between a transmitting and a receiving station, utilizing the space around the earth-atmosphere is called *space communication*. The earth's atmosphere plays a very interesting role in the propagation of e.m. waves from one place to another. This is possible due to change in air temperature, air density and electrical-conductivity. The absorption characteristics and height is also responsible for this like (i) the radiations in infrared region are absorbed by the atmosphere (ii) The ultraviolet radiations are absorbed by the ozone layer.

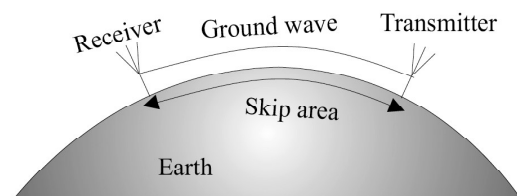
The following layer-types are considered to play main role in communication:

- *C-layer* at about 60 km above the surface of earth reflects e.m. waves in the frequency range 3 kHz – 300 kHz. It is used for direct long range communication.

- *D-layer* at a height of about 80 km reflects e.m. waves in the low frequency range (3kHz – 300kHz). It absorbs waves in the medium frequency range (300 kHz – 3MHz) and high frequency range (3 – 30MHz).
- *E-layer* is at a height of about 110 km. It helps in propagation of waves in the medium frequency range but reflects waves in the high frequency range in the day time.
- *F<sub>1</sub>-layer* at a height of about 180 km and lets most of the high frequency waves to pass through.
- *F<sub>2</sub>-layer* (at a height of 300 km in daytime and 350 km at night) reflects e.m. waves up to 30MHz and allows waves of higher frequencies to pass through.

You may recall from your earlier classes that, the atmospheric layer is close to the earth called the *troposphere* extends up to about 12 km above the sea level. The temperature in troposphere varies between 290 K (at the equator) to 220 K. Here, air density is maximum and electrical conductivity is minimum. The next layer up to about 50 km is called the *stratosphere*. An ozone layer is in the lower stratosphere extends from about 15 km to 30 km. The layer above the stratosphere and up to about 90 km is called the *mesosphere*. The minimum temperature at mesosphere is around 180 K. Beyond mesosphere up to 350 km, there is a zone of ionized molecules and electrons called the *ionosphere*. In ionosphere, temperature increases with height to about 1000 K. The ionosphere affects the propagation of radio waves. It is divided into D, E, F<sub>1</sub> and F<sub>2</sub> regions based on the number density of electrons, which increases with height from about 10<sup>9</sup> m<sup>-3</sup> in D-region to 10<sup>11</sup> m<sup>-3</sup> in region and 10<sup>12</sup> m<sup>-3</sup> in F<sub>2</sub> layer. These variations in temperature, density and conductivity arise due to different absorption of solar radiations at different heights and changes in composition etc.

The essential feature of space communication is that a signal emitted from an antenna of the transmitter has to reach the antenna of the receivers. Depending on the frequency (fo) of radio wave, the communication can occur as *ground wave*, *space wave*, *sky wave* and via satellite communication. A detailed aspect of this would be discussed in the following section.



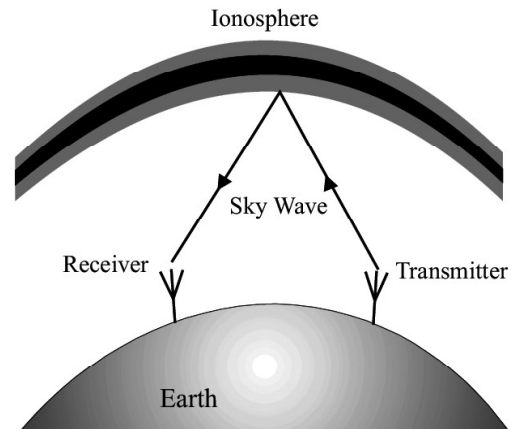
**Fig. 29.9 :** Ground wave propagation

- (i) **Ground Wave Propagation :** In ground wave propagation (Fig. 29.9), the electromagnetic waves travel along the surface of the earth. These can bend around the corners of the objects but are affected by terrain. A vertical antenna is used to transmit electromagnetic waves. If electric field *E* is vertical, and the magnetic field *B* is horizontal, then the direction of propagation *k* is horizontal and perpendicular to both *E* and *B* vectors. The material properties of the ground waves, such as its conductivity, refractive index and dielectric constant are seen to control propagation of such waves. That is why ground waves propagation is much better over sea than desert. In practice, ground waves are rapidly attenuated due to scattering by the curved surface of the earth. A larger wavelength results in smaller attenuation. That

is, ground waves are more useful as lower frequencies & constitute the only way to communicate into the ocean with submarines. Moreover, this mode of propagation is suitable for short range communication. For these reasons, ground wave propagation is used for radio wave (300 kHz – 3 MHz) transmission.

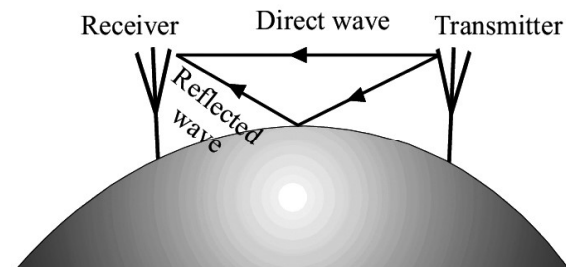
**(ii) Sky Wave or Ionospheric Propagation**

: In **sky wave** or **ionospheric propagation** (Fig. 29.10), the electromagnetic waves of frequencies between 3 MHz – 30 MHz launched by a transmitting antenna travel upwards, get reflected by the ionosphere and return to distant locations. In this mode, the reflecting ability of the ionosphere controls the propagation characteristics of the sky wave. The ionosphere acts as an invisible electromagnetic “mirror” surrounding the earth – at optical frequencies it is transparent, but at radio frequencies it reflects the electromagnetic radiation back to earth.



**Fig. 29.10** : Sky wave propagation

The maximum distance along the surface of the earth that can be reached by a single ionospheric reflection ranges between 2010 and 3000 km depending on the altitude of the reflecting layer. The communication delay encountered with a single reflection ranges between 6.8 and 10 ms. This mode of propagation is used for long- distance (short wave) communication in the frequency range approximately between 5 and 10 MHz. Above 10 MHz, the waves pass through the ionosphere and do not reflect back to the earth. It is, however, subject to erratic daily and seasonal changes due to variations in the number density and height of the ionized layers in the ionosphere. The composition of the ionosphere at night is different than during the day because of the presence or absence of the sun. That is why international broadcast is done at night because the reflection characteristics of the ionosphere are better at that time.



**Fig. 29.11** : Space-wave propagation

**(iii) Space Wave Propagation** : You

might have seen very high antennas at radio station. These are used for broadcasting. In space wave propagation (Fig. 29.11), some of the VHF radio waves (30 MHz – 300MHz) radiated by an antenna can reach the receiver travelling either directly through space or after reflection by the curvature of the earth. (Note that earth reflected waves are different from ground waves. In practice, direct wave mode is more dominant. However, it is limited to the so-called *line-of-sight* transmission distances and curvature of earth as well as height of antenna restricts the extent of coverage.



- (iv) **Satellite Communication** : So far you have learnt that ground waves suffer conduction losses, space waves have limitations due to line of sight and sky waves penetrate the ionization beyond a certain frequency. Some of these difficulties were circumvented with the launch of communication satellites in the 1950s. Satellite communication has brought about revolutionary changes in the form and format of transmission and communication.

The basic principle of satellite communication is shown in Fig. 29.12. The modulated carrier waves are beamed by a transmitter directly towards the satellite. The satellite receiver amplifies the received signal and retransmits it to earth at a different frequency to avoid interference. These stages are called up-linking and down-linking.

As we have seen already in connection with communication with light waves, the capacity of a communication channel can be increased by increasing the frequency of communication. How high up can we go in frequency? You now know that the ionosphere does not reflect waves of frequencies above 10 MHz, and for such high frequencies we prefer space wave propagation with direct transmission from tall towers. But this line-of-sight transmission also has a limited range or reach. Hence for long-range wireless communication with frequencies above 30 MHz, such as for TV transmission in the range of 50 - 1000 MHz, communication through a satellite is used.

The gravitational force between the earth and the satellite serves as the centripetal force needed to make the satellite circle the earth in a *free-fall* motion at a height of about 36,000 km. An orbit in which the time of one revolution about the equator exactly matches the earth's rotation time of one day is called a *geo-stationary* orbit, i.e., the satellite appears to be stationary relative to the earth. Ground stations transmit to orbiting satellites that amplify the signal and retransmit it back to the earth. If the satellites were not in geo-stationary orbits, their motion across the sky would have required us to adjust receiver antenna continually.

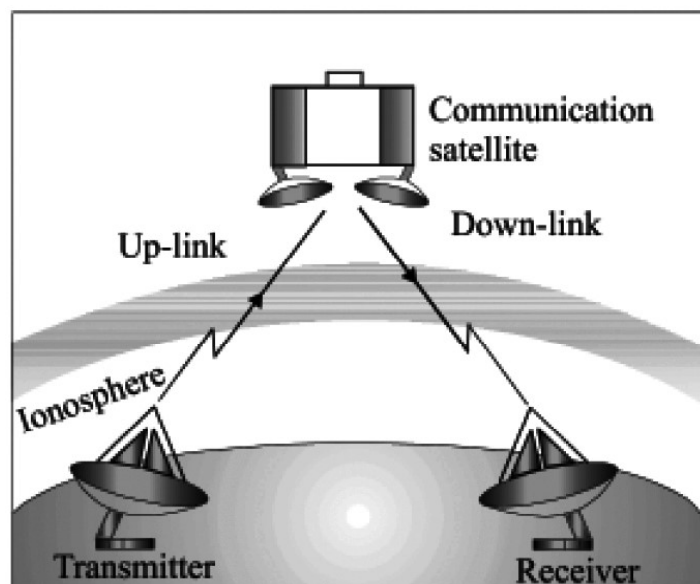


Fig. 29.12 : Satellite communication.

Two other orbits are also currently being used for communication satellites: (i) *polar circular orbit* at a height of about 1000 km almost passing over the poles (i.e., with an inclination of  $90^\circ$ ), and (ii) *highly elliptical inclined orbit* (with an inclination of  $63^\circ$ ) for communications in regions of high altitudes.

### Intext Questions 29.3

1. Choose the correct option in each case:
  - (a) Modulation is used to
    - (i) reduce the bandwidth used
    - (ii) separate the transmissions of different users
    - (iii) ensure that information may be transmitted to long distances
    - (iv) allow the use of practical antennas.
  - (b) AM is used for broadcasting because
    - (i) it is more noise immune than other modulation systems
    - (ii) it requires less transmitting power compared to other systems
    - (iii) it avoids receiver complexity
    - (iv) no other modulation system can provide the necessary bandwidth for faithful transmission.
2. Why do you hear some radio stations better at night than in the day?
3. Choose the correct option in each case:
  - (a) Frequencies in the UHF range normally propagate by means of
    - (i) Ground Waves
    - (ii) Sky Waves
    - (iii) Surface Waves
    - (iv) Space Waves
  - (b) Satellites are used for communication
    - (i) With low (<30 MHz) frequencies and for a small range
    - (ii) With low (<30 MHz) frequencies and for a long range
    - (iii) With high (> 30 MHz) frequencies and for a small range
    - (iv) With high (>30 MHz) frequencies and for a long range

### EDUSAT

*The Indian Space Research Organisation (ISRO), Department of Space, Government of India, launched an exclusive education satellite EDUSAT in Sept. 2004. The satellite has its footprints all over the country and operates in  $K_U$  (the word K in German means kurz - unten; below NATO-K band; 2240 GHz) band. It is designed to provide services for seven years. This satellite has capability for radio and TV broadcast, Internet-based education, data broadcasting, talk-back option, audio-video interaction, voice chat on Internet and video conferencing. It has opened up numerous possibilities: a teacher of a leading educational*

institution in a city may video-conference with students of a remote school, or school drop-outs in villages may receive Internet-based education support and get back into mainstream education system. EDUSAT has the capability of telecasting 72 channels. A large number of networks have been created by state governments and national institutions including NIOS. Such networks are being successfully used to impart education even in regional languages.

### Additional information on common communication devices

So far we have discussed the basic physics of signal transmission and detection process. Now we concentrate few typical electronic devices used in communication.

**Radio :** In radio transmission, a high-frequency carrier wave is first modulated by the information signal (voice or music). The modulated carrier is then transmitted in space by an antenna. The instrument used to detect such a modulated carrier is commonly called as a radio or a transistor. A radio **receiver** intercepts the radio waves with an antenna or aerial. Antenna selects the desired signal with the help of tuned LC circuits and amplifies the weak radio-frequency (rf) signal. After that, it decodes the audio signal from the radio wave through the process of demodulation and amplifies the audio signal. The amplified audio signal is then fed to a loudspeaker which reproduces the input audio signal. The block diagram of radio receiver circuit is shown in the Fig. 29.13.

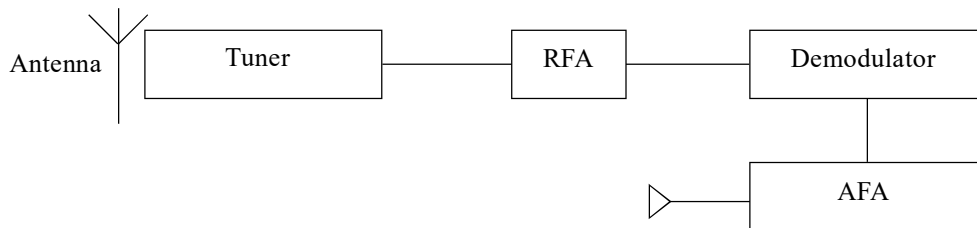


Fig. 29.13 : Radio receiver circuit

**Fax :** A facsimile or fax is seen in every modern office. When connected to a telephone line, a fax allows us to communicate our message to someone else far away. It is almost instant, faster than any other communication in delivery of documents. It can transmit any document such as printed material, hand written copy, picture or diagram and reproduce it faithfully.

Fax machines have existed for more than 150 years. However, its design has continuously evolved in accordance with developments in electronic circuitry. Present day fax machines are very fast. Basically a fax machine consists of a scanner and a printer. The scanner scans the document electrically, line by line, and transmits it by telephone. The message is decoded at the other end and reproduced through a printer. That is, the scanner and the printed are not co-located; these are linked by a telecom medium. To scan a document, light from a powerful source is focused and reflection of each small element on a line is evaluated progressively from left to right and it is repeated line by line till the end.

**Modem :** Modems came into existence in 1960's as a way to connect terminals to computers over the phone lines. The word **modem** is coined out of the words modulator

**demodulator.** A series modem modulates the digital data from a computer into an analog signal compatible with an analog communication channel (telephone line). On the other hand, receiving modem demodulates the analog signal into a digital data for the receiving computer. It means that the same modem works as a modulator when sending messages and as a demodulator when receiving messages. Wireless modems convert digital data into radio signals and back.

## WHAT YOU HAVE LEARNT

- *The essential elements of a communication system are (i) a transmitter (ii) a medium or mechanism to carry the signal over long distances, and (iii) a receiver to intercept the signal and retrieve the information.*
- *An antenna or aerial is essentially a system of conductors, which is an effective radiator and absorber of electromagnetic waves in the desired radio frequency region.*
- *Analog signals are physical signals that vary continuously with time while digital signals have the form of discrete pulses.*
- *Digital communication systems are more efficient, give better performance, and greater flexibility than their analog counterparts*
- *AM radio is broadcast on three bands, the Long wave at 144 – 351 kHz (in the LF), the Medium wave at 530 – 1,700 kHz (in the MF), and the Short wave at 3 – 30 MHz (HF).*
- *An analogue signal is completely described by its samples, taken at equal time intervals  $T_s$ , if and only if the sampling frequency  $f_s = 1/T_s$  is at least twice the maximum frequency component of the analogue signal.*
- *Low frequencies cannot be transmitted to long distances using aerials or antennas of practical dimensions. Low-frequency messages are loaded on a high frequency carrier signal by a process called modulation. In amplitude modulation (AM), the amplitude of a high-frequency carrier wave are modified in accordance with the strength of a low-frequency information signal. In frequency modulation (FM), the amplitude of the carrier wave remains constant, but its frequency is continuously varied in accordance with the instantaneous amplitude of the information signal, i.e., the frequency of the modulated carrier wave varies from a minimum to a maximum corresponding to the minimum and maximum values of the modulating signal.*
- *In the digital pulse code modulation (PCM) technique, the modulating signal is first sampled, the magnitude (with respect to a fixed reference) of amplitude of each sample is quantized, and then the binary code is usually transmitted modulating an analogue current in a landline.*
- *Electrical communications channels are wireline (using guided media) or wireless (using unguided media).*
- *Multiplexing refers to the process of simultaneous transmission of different messages (each with some frequency bandwidth) over the same path way. The higher the frequency of the carrier, the higher is its message-carrying capacity.*

- Comparing the different wireline channels, the communication capacity of visible light (of frequency of about  $10^{14}$  Hz) in an optical fibre is thus much larger than that of typical microwave (of frequency of about 109 Hz) in a metallic conductor.
- In the wireless radio transmission, a system of conductors called antenna or aerial launches the carrier radio waves in space and also detects them at the receiver location. The propagation of radio waves in the atmosphere depends on the frequency of the waves. Low and medium frequency radio waves up to about 1 MHz are used in ground (or surface) wave communication. Medium frequency (MF) waves of 300 kHz – 3 MHz are largely absorbed by the ionosphere. The high-frequency (HF) waves of 3 – 30 MHz are, however, reflected back by the ionosphere. VHF and UHF waves are transmitted either by direct line-of-sight using tall towers (space wave or tropospheric propagation), or by beaming to artificial satellites and broadcasting from there.
- The cellular or mobile telephone system works as a radio network in which a city is divided into ‘cells’ of 1 km to 50 km in radius, and each cell is covered by one base station. A cellular phone contains a low-power transmitter and a low-power receiver.
- *In a typical modern-day communication system, the information is in the form of electrical signals (voltage or current).*

### TERMINAL EXERCISE

1. What are the essential elements of a communication system?
2. What is an antenna?
3. What are the important characteristics of a receiver in a communication system?
4. Distinguish between the terms analogue and digital signals. Define a ‘bit’.
5. The VHF band covers the radio frequency range of 30 – 300 MHz. Using the known relationship of speed to frequency and wavelength of an electromagnetic wave, determine the VHF wavelength range in vacuum. Take the speed of light in vacuum to be  $3 \times 10^8 \text{ ms}^{-1}$
6. Long distance radio broadcasts use shortwave bands. Explain.
7. Satellites are used for long distance TV transmission. Justify.
8. Explain sampling theorem.
9. What do you mean by modulation? Explain.
10. Explain the process of demodulation.

## ANSWERS TO INTEXT QUESTIONS

### 29.1

1. 30 kHz - 300 MHz
2. The maximum size of antenna is related to the frequency of the signal. It is designed to transmit or receive and capture maximum radial part.

### 29.2

1. em waves consist of oscillatory electric and magnetic fields. It carries momentum and radiates energy.
2. 
$$\lambda = \frac{c}{\nu} = \left( \frac{3 \times 10^8 \text{ ms}^{-1}}{30 \times 10^6 \text{ s}^{-1}} \right) = 10 \text{ m}$$
3.  $10^{14}$  to  $10^{15}$  Hz

### 29.3

1. (a) iv (b) (iii)
2. The composition of ionosphere and reflection characteristics are good at night (absence of sun).
3. (a) iv (b) (iv)

## ANSWERS TO TERMINAL EXERCISE

2. 10 m - 1 m

